Techniques de classification pour l'identification et la prédiction non intrusive de l'état des charges dans le bâtiment

Kaustav Basu

UNIVERSITÉ DE GRENOBLE

THÈSE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Génie Électric**

Arrêté ministériel : 7 août 2006

Présentée par
M. Kaustav BASU

Thèse dirigée par
M. Seddik BACHA et **M. Vincent DEBUSSCHERE**

préparée au sein du
laboratoire de Génie Electrique de Grenoble (G2Elab)
dans **l'école doctorale d'Électronique, Électrotechnique, Automatique et Traitement du signal (EEATS)**

Techniques de classication pour l'identification et la prédiction non intrusive de l'état des charges dans le bâtiment.

Thèse soutenue publiquement le **14 novembre 2014**,
devant le jury composé de :

Cristian NICHITA, Président
Professeur, Université du Havre
Bruno SARENI, Rapporteur
Professeur, INP-ENSEEIHT/LAPLACE, Toulouse
Dirk BENYOUCEF, Rapporteur
Professeur, Université de Furtwangen, Allemagne
Hamid BEN AHMED, Examinateur
Maître de conférences HDR, Ecole Normale Supérieure de Rennes
Ahlame DOUZAL, Examinatrice
Maître de conférences HDR, Université Joseph Fourier Grenoble
Seddik BACHA, Directeur de thèse
Professeur, Université Joseph Fourier Grenoble
Vincent DEBUSSCHERE, Encadrant de thèse
Maître de conférences, Insitut National Polytechnique de Grenoble

UNIVERSITY OF GRENOBLE

Doctoral school EEATS

(Électronique, Électrotechnique, Automatique et Traitement du Signal)

T H E S I S

for obtaining the title of

Doctor of Science

of the University of Grenoble

Mention GE : Génie Électrique

Presented by

Mr Kaustav Basu

Classification techniques for Non-intrusive load monitoring and prediction of residential loads

Thesis supervised by

Mr Seddik Bacha and Mr Vincent Debusschere

prepared at the

Grenoble Electrical Engineering Laboratory (G2ELab)

presented on November 14^{th}, 2014

Jury :

President :	Cristian NICHITA	-	University of Le Havre
Reviewers :	Bruno SARENI	-	INP-ENSEEIHT/LAPLACE, Toulouse
	Dirk BENYOUCEF	-	Furtwangen University, Germany
Examiners :	Hamid BEN AHMED	-	Ecole Normale Supérieure, Rennes
	Ahlame DOUZAL	-	University Joseph Fourier, Grenoble
Supervisor :	Seddik BACHA	-	University Joseph Fourier, Grenoble
Co-supervisor :	Vincent DEBUSSCHERE	-	Grenoble Institute of Technology

Table des matières

General Introduction **1**

I State of the Art **5**

1 Non-Intrusive Appliance Load Monitoring **7**

1.1 Problem statement 8

1.2 Non-intrusive load monitoring 8

1.3 Challenges 9

1.4 The French context 12

2 Appliance Future Usage Prediction **15**

2.1 Problematic 15

2.2 Predicting energy consumption 18

2.3 Energy Management from inside the buildings 19

II Qualitative Energy Data Analysis **23**

3 Database and Energy Consumption Analysis **25**

3.1 Data-set 25

3.2 Visualization Techniques 26

4 Clustering and Flexibility **35**

4.1 Categories of households 35

4.2 Categories of appliances 37

4.3 Load flexibility for the grids 39

III Non-Intrusive Load Monitoring **45**

5 Methodology of Temporal Classification **47**

5.1 Temporal sliding window 48

5.2 Meta-features 50

6 Classification algorithms **57**

6.1 Multi-label classification 58
6.2 The K-Nearest Neighbours classifier 62
6.3 Hidden Markov Model based classification 66
6.4 Imbalance in the data-set 67
6.5 Evaluation Strategies 68

7 Load Monitoring Results 71
7.1 Time-related observations 72
7.2 Energy variation between time steps 74
7.3 Identifying the loads 75
7.4 Load identification for defined categories of consumers 76
7.5 Sensitivity to the sampling rate 81
7.6 Comparison with standard non-intrusive load identification methods 82
7.7 Discussion 84

IV Appliance Usage Prediction 87

8 Methodology of Appliance Usage Prediction 89
8.1 Load forecasting algorithms 89
8.2 Problem statement 94
8.3 Overall proposed architecture 96
8.4 The oracle : for knowledge extraction 98
8.5 The predictor : for classification 101

9 Residential Load Prediction Results 105
9.1 The oracle features performance 105
9.2 Using different input categories and classifiers 107

10 Appliance Usage Prediction : using multi-label classifier 117
10.1 Context 118
10.2 Prediction the state of the loads in different conditions 120
10.3 Load prediction after load identification 123
10.4 Prediction algorithm and energy management systems 126

11 Case study : Residential Energy Management 129
11.1 Local Energy Management 130

11.2 The optimal usage of energy storage 132
11.3 Optimal Load Scheduling 138

Conclusion and Perspectives **147**

List of figures **149**

List of tables **153**

Table of Abbreviations and Acronyms **157**

French Abstract **159**

11.4 Contexte des travaux 159
11.5 Classification des charges 161
11.6 Résultats 167
11.7 Conclusion 171

Bibliographie **173**

A Annexure **179**

A.1 Sample dataset 179
A.2 Database groping (parallel coordinate plot) 179
A.3 Load identification frequency analysis 181

General Introduction

Context

There are currently two main ways to deal with the worldwide growing energy consumption. The first possibility is to globally increase energy production capabilities. The second is to head toward a more efficient energy consumption. Considering the second alternative, we decided to work on energy consumption of one of the elementary loads on the grids : buildings.

According to [Observatoire Electricité 2014], the residential sector represents 44 % of the total energy consumption in France with 69 Mtoe consumption from a total of 154 Mtoe. The residential sector represents an important factor in the French energy consumption and can be associated with substantial savings in terms of energy and/or money. For information, the transport sector is the second largest with 32 % of the total energy consumption in France, then the industry with 21 % and the agriculture with 3 %.

Focusing on the buildings sector, in the near future, the main issue concerns civil engineering and the thermal insulation of buildings. But in the long term, issues concern local integration of renewable energy and smarter buildings connected to smarter grids [Clastres 2010]. A relevant knowledge of appliances consumption in buildings is needed in order to better control or monitor energy consumption.

These controls and monitoring have two main goals. The first is to decrease the energy consumption of buildings and/or decrease the electricity bill of inhabitants. The second is to propose more tools to the grid managers in order to better manage an increasing intermittent energy production due to the increasing renewable energy integration in the grids.

Therefore, monitoring and control of appliances consumption has two different purposes whose business models are still in development. First, from the point of view of a *user or inhabitant*, having information on appliances level usage can lead to reduced cost through a reduction of the energy consumption or possible ancillary services (unbalancing requests, load shading or energy price variations, etc.). Second, from a *smart grid or grid manager* point of view the control of more loads represent more possibility of actions for maintaining stability of the grids, i.e. more flexibility and reliability (reduce peak demand by eliminating electricity use, or by shifting it to non-peak times, etc.). These services represent elementary bricks of an energy management system whose privacy limits has still to be defined. They are explored in the work, firstly, by *load identification* and secondly, by *prediction of load energy consumption*.

Another aspect for energy management is in terms of smart meters. According to the latest energy policy, the objectives in France is to reach the target of 90 % of smart power meters penetration by 2020. To achieve this goal new data analysis mechanisms have to be proposed to inhabitants for their satisfaction and energy costs reduction. Just a transfer from an analog to a digital system is not good enough for the customers. A comprehensive and qualitative data analysis mechanism has to be proposed coupled with the subsequent load management strategies that have to evolve. The challenges in energy monitoring and management is well defined by Mickael MacKenzie, Vice President of digital energy services at Schneider Electric :

"The difficulty in developing a core energy management strategy is a lack of visibility into the full energy story of a facility. Getting insight into how energy is being consumed, what normalizing factors create inefficiencies and how wastes, such as emissions or effluents are managed, is critical to understanding how human behavior has an impact.

Data is frequently collected using utility bills and spreadsheets making it difficult to correlate energy events with production events inside a plant. The proper data capture strategy can turn raw data into information into wisdom, enabling continuous improvement strategies".

Contribution

The major contributions in the work are as follows :

- A relevant pixel based energy data-visualization approach is implemented and its parametric intricacies discussed. Modern visualization tools allow to move beyond daily, monthly and yearly curves and visualize various patterns exhibited in the energy consumption data (Chapter 3).
- The work propose a generic temporal classification approach after summarizing approaches already available in the non-intrusive load monitoring domain. The novelty of the proposed approach lies in the fact that it is applicable to the current smart meter (especially considering their sampling rates) and that it reduces privacy concerns by primarily concentrating on high energy consuming appliances (Chapter 5).
- A novel multi-label classification approach is proposed and a selection of algorithms implemented for comparison. Temporal distance based approach and sequence learning technique are compared (Chapters 6 and 7).
- A novel and generic appliance future usage prediction is proposed after summarizing other techniques already available in the load prediction domain. The prediction algorithm is made to be integrated in a three layer software architecture defined for smart home monitoring and control. The requirements of an appliance prediction system is highlighted (Chapter 8).
- An expert knowledge based model is compared for standard classification techniques and different aspects of the model are discussed. A real time implementation of the appliance prediction system is also evaluated and performance on selected appliances are compared (Chapters 9 and 10).

These contributions are highlighted in the following papers :

1. K. Basu, V. Debusschere, S. Bacha, *"Non Intrusive Load Monitoring : A Temporal Multi-Label Classification Approach", IEEE transaction on Industrial Informatics, accepted*[Basu 2014].
2. K. Basu et al *"A prediction system for home appliance usage"*, Elsevier Journal for Energy and Buildings, Energy and Buildings, 2013, vol. 67, p. 668–679.
3. K. Basu, V. Debusschere, S. Bacha, *"A prediction system for home appliance usage"*, Industrial Electronics Society, IECON 2013-39th Annual Conference of the IEEE, 2013.
4. K. Basu, V. Debusschere, S. Bacha, *"Load identification from power recordings at meter panel in residential households"*, IEEE XXth International Conference on Electrical Machines (ICEM), 2012, [Basu 2012b].

5. K. Basu, V. Debusschere, S. Bacha, *“Appliance usage prediction using a time series based classification approach”*, IECON 2012-38th Annual Conference of the IEEE, 2012.
6. K. Basu, M. Guillame-Bert, H. Joumaa, S. Ploix, J. Crowley, *“Predicting Home Service Demands From Appliance Usage Data”*, International Conference on Information and Communication Technologies and Applications (ICTA). Jointly with the 17th International Conference on Information Systems Analysis and Synthesis (ISAS), 2011, Orlando, Florida, USA. [Basu 2011]
7. V. Debusschere, K. Basu, S. Bacha, *“Identification et prédiction non intrusive de l'état des charges dans les bâtiments résidentiels à partir de mesures compteur à échantillonnage réduit”*, Symposium du Génie Electrique, 2014, Cachan, France. [Debusschere 2014a]
8. V. Debusschere, W.R.L. Garcia, K. Basu, S. Bacha, *“Système de management énergétique résidentiel prédictif sous critères technico-économiques"*, Symposium du Génie Electrique, 2014, Cachan, France. [Debusschere 2014b, Debusschere 2014c]
9. V. Debusschere, W.R.L. Garcia, K. Basu, S. Bacha, *“Bilan sur cycle de vie des flux énergétiques dans les bâtiments résidentiels incluant de la production et du stockage”*, 3ème Conférence francophone sur l'éco-conception en Génie Electrique CONFREGE, 2014, Albi, France. [Debusschere 2014d]

Summary

Energy management for residential homes and/or offices requires both identification of the load inside the buildings from the power meter and prediction of the future usages or service requests of these appliances. The aim of the work is to identify residential appliances from aggregate power readings at the power meter and to predict their states in order to manage and possibly to minimize their energy consumption. For this purpose, our work is divided in two distinct modules : *Appliance identifcation* and *future usage prediction*. Both identification and prediction are based on multi-label learners which takes inter-appliance co-relation into account.

The *residential buildings* sector is mainly considered in this work because of the significant variability of load profiles, but the tools and methodology are totally applicable to any other kind of buildings.

1. The first part of the work concerns the identification of electrical appliances usages from the smart meter monitoring. The main objective is to be able to identify individual loads from the aggregate power consumption in a non-intrusive manner. High energy consuming appliances are identified at low sampling rate using novel set of meta-features for this domain.
2. The second part concerns future usage prediction. A generic model for future usage of appliances is presented and different strategies are discussed.

Finally, this work is based on a real residential dataset of 100 houses monitored every 10 minutes during one year (including weather information).

Première partie

State of the Art

Chapitre 1

Non-Intrusive Appliance Load Monitoring

1.1 Problem statement **8**
1.2 Non-intrusive load monitoring **8**
1.3 Challenges **9**
1.3.1 Works on load monitoring 10
1.3.1.1 High sampling rate Nialm 11
1.3.1.2 Low sampling rate Nialm 11
1.4 The French context **12**
1.4.1 Our contribution 13

Introduction

From the *inhabitants* point of view, an energy management system applied to loads in buildings lets customers adjust their energy consumption according to an expected level of comfort, energy prices variations and sometimes environmental impacts (for example CO_2 equivalent emissions). Such demand side management strategies need an accurate evaluation of the amount of energy that can be controlled and through which loads. Therefore, identifying the usage of each appliance is one of the core issues in the field of smart buildings energy management.

From the *smart grid* point of view, receiving information on the usages of appliances (especially deferrable loads) helps to manage the energy distribution [Strasser 2013, Palensky 2011], especially for the integration of more fluctuating energy sources (i.e. renewable). The energy management depends on appliances : some can be postponed (washing machine, etc.) and some cannot be (television set, etc.). In this field, there already exist strategies defined as *demand response* [Siano 2014] to reduce peak demand by eliminating electricity use, or by shifting it to non-peak times. The proper use of these techniques can depend on *time of use pricing*, then on energy prices variations and ultimately on consumer acceptance. From the point of view of energy providers, load identification can also play an important part in future prediction of usages of particular appliances [Basu 2013b] where the process of historical data collection is made as less intrusive as possible.

At the moment, current power meters only report whole-residence data. It is required to separate and subsequently identify the total load into its constituent components, i.e. appliances as shown in the Figure 1.1. In order to avoid indirect disaggregation, the appliances within the house could be monitored directly, but at the costs of manufacturing and installing many new devices in the houses, inconvenience to the user and the fact that new sensors have to be installed for any new appliances. Non-intrusive methods propose an attractive alternative with reduced cost and manual overheads.

1.1 Problem statement

Non-intrusive load monitoring deals with the disaggregation of individual appliances from the total load at the smart meter. So if a load curve L monitored at a power meter is the sum of three loads consuming respectively L_1, L_2 and L_3, then the the task is to determine the state of L_1, L_2 and L_3 individually with the only knowledge of L. This task is illustrated in the Figure 1.2.

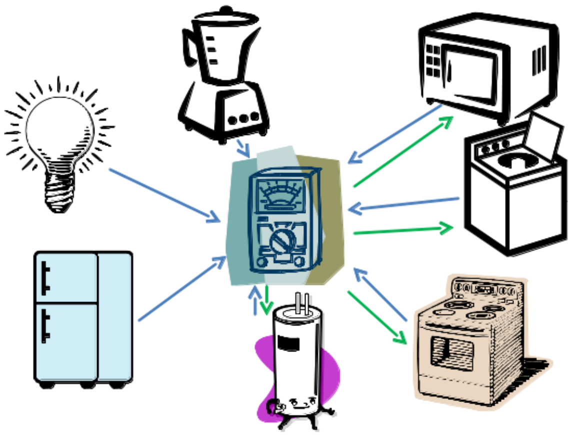


FIGURE 1.1 – Identification and management of appliances through the power meter

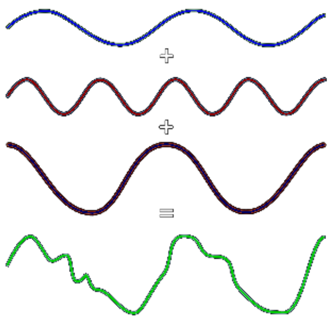


FIGURE 1.2 – The principle of signal separation

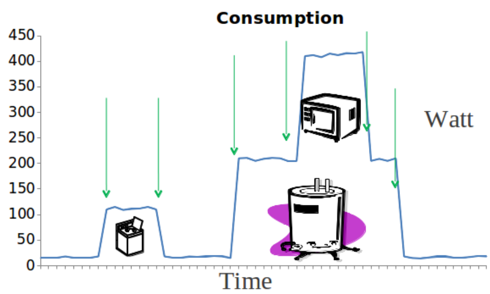
1.2 Non-intrusive load monitoring

The smart meters are one of the fundamental units of the smart grids, as many further applications depend on the availability of fine-grained information on energy consumption and production. From the energy consumer's point of view, the access to fine-grained data can be more than just load curves and energy readings, but also consumption information and potential reduction. As a simple example, modern visualization tools can easily show comparisons to previous days and weeks. These visualization give a detailed picture of the energy consumption of a building down to the identification of individual devices (e.g., washing machines, water heater), as each device has its typical load curve.

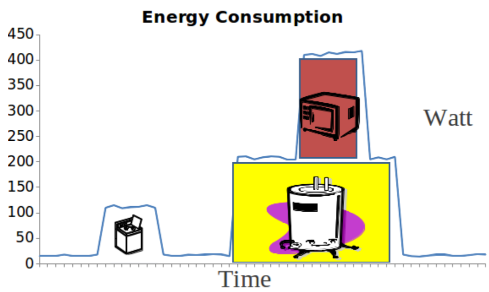
From the service provider's point of view, comparisons and useful information can be provided to peer groups of consumers having similar consumption patterns (in terms of size and appliances used). In addition, devices can be identified from their load profile, which can be visualized additionally in a non-intrusive mechanism [Farinaccio 1999]. The similitude of consumption for specific appliances in a group of houses is significant information for grid managers who wish to use consumption flexibility at a proper level.

Load monitoring instrumentation used to involve complex data-gathering hardware but more simple software mechanisms. All the appliances of interest were directly monitored using wires or power-line carrier techniques or radio signalling connected to a central data-gathering unit. Conversely, a Non-Intrusive Appliance Load Monitoring (NIALM) or Non-Intrusive Load Monitoring (NILM) system consists of a simpler hardware part and a more complicated software mechanism.

Load separation methods can be classified based on the intrusiveness of the training process and the nature of the classification algorithm (event-based or non-event-based). Event-based algorithm tries to detect ON/OFF transitions whereas non-event-based methods tries to detect whether an appliance is ON during the whole sampled duration. Figure 1.3 proposes the two cases, event-based in Figure 1.3a and non-event-based in Figure 1.3b.



(a) Event-based load monitoring method



(b) Non-event-based load monitoring method

FIGURE 1.3 – Two types of methods for non-intrusive load monitoring

1.3 Challenges

Researchers have been working on the NIALM problem for the last two decades. The challenges to a NIALM problem is both technical and social.

Technically, it is built on top of the premise that the study of how the variation over time of the global energy consumption of a building can lead to information about the appliances that have advocated these changes. Most of the approaches were based on signal processing at a high sampling rate (1 second typically) to evaluate the appliance load signature and subsequently to use pattern recognition techniques for identification from previously trained classifiers. This requires the installation of a sensor for each appliance within the house and is then naturally restricted by this important system of load monitoring (without speaking about the cost). Also, the cost benefit for the user has to be carefully analyzed before developing this kind of solution. The appliance usage being different from one user to the other, the variability of consumption patterns is not compatible with a systematic benefit.

Socially, a major hurdle in the NIALM research is the privacy concerns of the user as the appliance usage can be co-related with user behavior. For example, the time at which the lights are shut d own can be assumed to be the sleeping time of the inhabitant.

Finally, a non-intrusive method which works for all the range of appliances within the house is yet to be developed.

1.3.1 Works on load monitoring

Traditionally the NIALM consists of the six overlapping data flow phases [Birt 2012]. The first is *data acquisition* followed by *data processing*, *event detection*, *feature extraction*, *event classification* and finally *energy computation* as shown in the Figure 1.4.

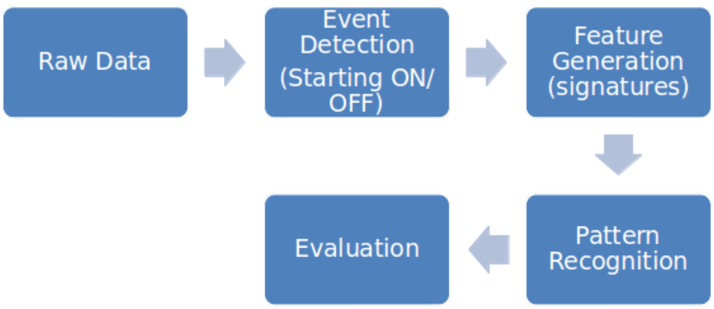


FIGURE 1.4 – Non-Intrusive Load Monitoring work flow

The pioneering work in load separation was started by Hart [Hart 1992] in the beginning of the 90's. The methods were proposed to identify individual appliances from their ON/OFF transitions. Appliance transitions result in corresponding changes in the overall power consumption monitored at the power meter. This pioneering signal processing technique is shown in Figure 1.5.

Methods were then proposed to identify individual appliances from their ON/OFF transitions. From that time, most of the approaches were event-based and at a high sampling rate, typically less than one second. This rate is required as event-based methods depends on state switching detection. The sampling rate is defined compared to the variation of the state of the loads. Events recognition works well for high sampling rates but fails most of the time at low sampling rate.

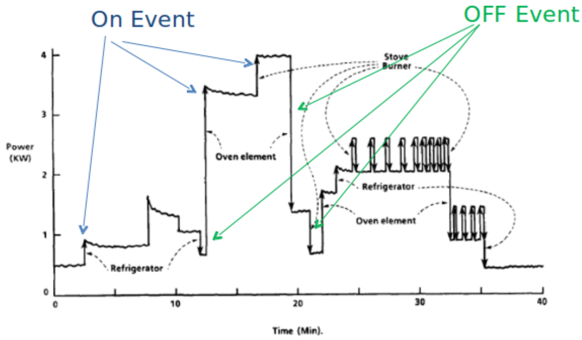


FIGURE 1.5 – Non-Intrusive Load Monitoring [Hart 1992]

1.3.1.1 High sampling rate NIALM

In the last two decades, there have been considerable amount of work to this effect. Each new method proposes to reduce the limitations of the previous ones both in term of signatures or applying state of the art pattern recognition techniques. The identified features are known as appliances signatures. Approaches typically consist of identifying the steady state or in some cases transient state features [Zeifman 2011, Li 2013]. Subsequently, these signatures are matched with earlier learned models using a pattern recognition algorithm [Berges 2010, Bier 2013]. The drawbacks of these approaches are mainly hardware requirement due to high sampling rates and the impracticality of the process being totally non-intrusive [Fernandes 2013, Norford 1996].

These methods do not fit well into the smart meter sampling rate, so separate device has to be installed for training, visualization and communication to the grid. This is a major drawback for these methods, commercially and practically speaking. The load separation at a high sampling rate of all the appliances also raise privacy concerns as user activity can be easily detected, interpreted and monitored [Birt 2012].

1.3.1.2 Low sampling rate NIALM

At a low sampling rate, switching events are difficult to detect so event-based methods are more suited. The major issue at low sampling rate is that low energy consuming devices are difficult to be detected. However, high energy consuming appliances, such as water heater or washing machine can still be identified with reasonable precision even at sampling rate of 15 minutes for example [Kalogridis 2010, Prudenzi 2002].

Considering the constraint of low sampling rate, the differentiation of the methods is directly dependent on the choice of algorithms. Some algorithms have already been implemented and tested in the field of load monitoring.

A method partially disaggregating total household electricity usage into five load categories has been proposed at a low sampling rate in [Kolter 2010] where different sparse coding algorithms are compared and a Discriminative Disaggregation Sparse Coding (DDSC) algorithm is tested. A feature-based Support Vector Machine (SVM) classifier accuracy is also mentioned but is not presented. The method of [Kolter 2010] is an implementation of the blind source separation problem, which aims at disaggregating mixture of sources into its individual sources. A classic example for this would be the problem of identifying individual speakers in a room having multiple mikes placed at different locations. In the NIALM context, the problem is undermined as there is only one mixture and a large number of sources. Another issue using blind source separation is the assumption of no prior information about the sources. On the contrary, in the NIALM context, the sources (appliances) do have separate usage patterns which could be used. Nevertheless, blind source separation still remains a promising direction of research in this domain.

Temporal graphical models such as Hidden Markov Models (HMM) also have been promisingly used in this domain as they are a classical method for sequence learning [Parson 2011]. They have been successfully used in many domains, especially speech recognition. In the NIALM context, the problem is to learn the model parameters given the set of observations as input sequence and appliances states as output. HMM also considers sequential patterns in consumption but in the NIALM context, at a very low sampling rate it seems to have a sensibility to training noise.

Among others, the algorithms SVM and HMM have a more detailed presented in Chapter 6. They also have been implemented and used in this works.

1.4 The French context

In 2010, the French distribution system operator "Electricité Réseau Distribution de France", ERDF, launched an Automated Metering Management project (AMM) that aims to implement 300 000 smart meters in France [1]. The smart meters, however, present low sampling rates starting from 10 minutes to one hour. These low rates of sampling considerably reduce the hardware complexities of the process. As an opportunity, most of the high energy consuming appliances have low frequency of usage, typically once a day. The identification of these appliances is useful, and can still be performed under restrictive conditions on the sampling rate. However, methods proven to be effective with a high sampling rate may not be as effective with a low sampling rate due to the difficulty to detect the appliance signatures.

A situation is considered where a user gives a recording (time stamped) of the usage of his high energy consuming appliances for a week or two (through e.g. a smart phone application) and subsequently gets his energy management plan for the year. There is no need of any particular power recordings other than the one of the household power meter. In cases where the users cannot monitor the usage of the appliances, inexpensive ON/OFF sensors can be used for the training phase only. These sensors have reduced privacy concerns ; they are only monitoring high consuming appliances and thus are more acceptable to users who don't desire their own behavior to be monitored.

These appliances may be, finally, controlled by a local energy management system (private or aggregating many consumers) responding to regional grid manager *flexibility requests* (through automatically shutting downs, shifting or shading the loads). In order for the inhabitants to let that happen, economic incentives will have to be proposed (real time pricing, financial compensation, etc.) and are still to be developed.

1. `http://www.erdfdistribution.fr/EN_Linky`

There are also other benefits of using the above mentioned technique. In many modern smart meter applications the electricity usage can be visualized by the user. He gets his electricity usage information but this one will lack qualitative information regarding what appliances are responsible for the total load and to what extent. Knowing the consumption of a particular appliance will give incentives to the user on its future usage.

An illustrative example that can be seen is in the super-market bill though the breakdown of the total bill into individual purchases. The benefit to the client is trivial (he knows what costs the most and possibly won't buy it anymore) but the supermarket chain also gets valuable information about the client purchase patterns. This has many benefits in terms of advertisements of products through client specific promotions.

1.4.1 Our contribution

The context of the training process offers two possibilities. The first one is to use supervised machine learning methods with the assumption of the availability of prior data for training. This training mechanism drastically increases the complexity of the monitoring process both in terms of cost and time. The second possibility is to use unsupervised disaggregation methods where no prior training data is used. This requires a post-processing phase where the appliances are labeled manually.

One of the strengths of the proposed method is to be a non-event-based approach with a very short and non-intrusive training period. In this work, a novel data collection mechanism is suggested and can be practically implemented due to of its procedural simplicity. Furthermore, an appliance state detection mechanism is presented and compared with the traditional NIALM mechanism. We consider 10 minutes to an hour to be a low sampling rate and less than a minute to be a high one [Basu 2012a].

Temporal classification using multi-label classification techniques presents an interesting alternative to signal analysis at low sampling rates and have not been tested in the field of load identification for households. We can divide into three parts the work on load identification presented in this manuscript.

- A way of visualizing data for identified appliances is addressed.
- Meta-features in the field of residential building appliances are proposed and a smart meter integrated methodology is formalized.
  - For the identification of the loads.
  - For the prediction of their future states.
- A variety of state of the art multi-label learners are applied to our dataset to find the most relevant learners in this field of research for different low sampling rates.

The limitations of these techniques are due to the sampling rate. In fact, with a very high sampling rate, privacy issues are raising and almost everything can be known about the users. Remaining at a low sampling rate is sufficient both for information to the user, for load management and scheduling (only high energy consuming appliances are interesting for the grid manager) and for the privacy concerns. Another limitation will be the variability of the efficiency of the methods depending on the appliances. These points will be addressed in the Chapter 7.

Summary

The chapter addresses the problem of load disaggregation at a low sampling rate which is compatible with the current smart meters. The different techniques used in the field are discussed. The techniques range from using probabilistic graphical models or event detection methods followed by pattern recognition. These techniques both have their own advantages and disadvantages which are discussed. The method proposed in this work is also introduced with a training procedure which also accounts for user consent. The advantages and limitations of the proposed method is also highlighted.

The global architecture of the identification process proposed in this work is shown in Fig. 1.6. All interactions between inhabitants, loads and the grid are centralized in one hub : *the classifier*. This classifier can be integrated or be an added part to a smart meter or an energy box. These components could directly interact with the local grid manager or be aggregated with other similar components in a "smart city" for example.

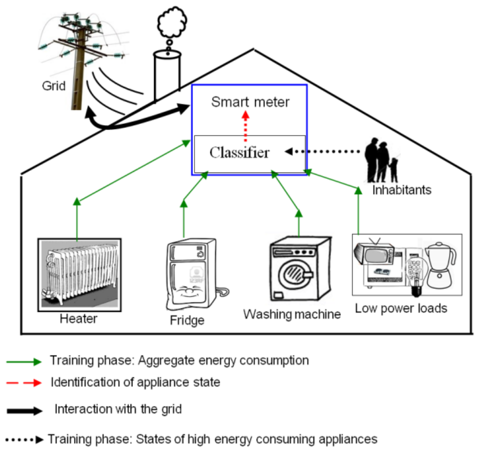


FIGURE 1.6 – Classifier architecture : Centralizing information of a smart household.

Chapitre 2

Appliance Future Usage Prediction

2.1 Problematic **15**

2.1.1 Residential Energy Management 15

2.1.2 Home Automation System 16

2.2 Predicting energy consumption **18**

2.3 Energy Management from inside the buildings **19**

2.1 Problematic

Reducing housing energy costs is a major challenge of the 21^{st} century. After thermal insulation of buildings, the issues are those of *local renewable energy integration* (solar, wind, etc) and *smart buildings*. In this last field, prediction of consumption is one of the key to a proper energy management system for loads in buildings.

2.1.1 Residential Energy Management

Load management allows inhabitants to adjust power consumption according to expected comfort, energy price variation, etc.. A home energy management system (Ems) is able to determine the best energy assignment plan and a good compromise between energy production and energy consumption [Ha 2006b].

In this work, the energy consumption is mainly restricted to the electricity consumption (local production and storage will be also considered as application). We consider a three-layers architecture system for the residential energy management system, consisting of an *anticipation* layer, a *reactive* layer and a *device* layer [Abras 2008]. This system is both able to satisfy the maximum available electrical power constraint and to optimize a compromise between user satisfaction and cost.

The objective of the anticipation layer is to compute plans for production and consumption of services. Uniqueness of housing systems involves a set of new issues in energy management : it is necessary to develop new tools and especially algorithms for globally optimized power management of the home appliances [Abras 2010, Elmahaiawy 2010, Ha 2010].

These algorithms should be able to anticipate difficult situations but also able to take into account the actual housing system state and the occupant expectations, without forgetting local production [Riffoneau 2010].

Anticipating problematic situations requires also prediction capabilities. Even if overall consumption is easier to predict, the usage of each appliance has to be known because of the dynamics of the demand side management.

Also, it is important to evaluate how much energy or money can be saved thanks to request to customers like unbalancing requests or energy price variations. The energy saving depends on appliances : some can be unbalanced, some can be postponed and some cannot be controlled. The overall concept of the smart home and its actors are described in Figure 2.1.

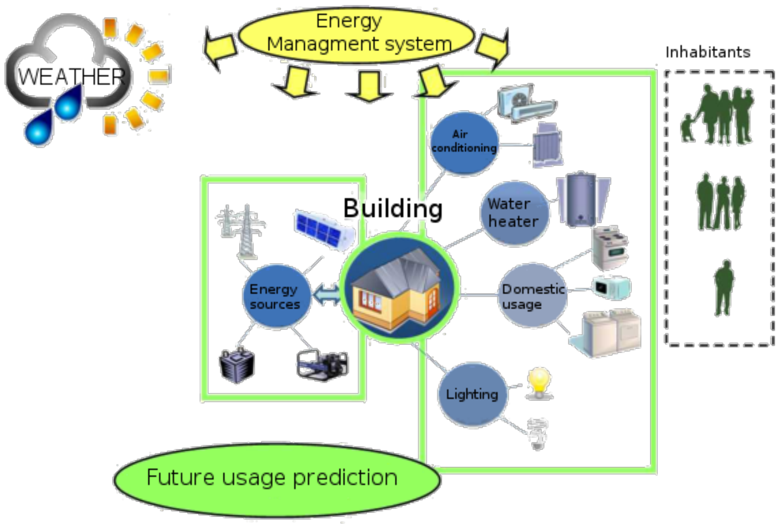


FIGURE 2.1 – Actors involved in the Smart Home concept

Based on the information obtained in the learning process, the prediction algorithm needs to be adaptive in order to forecast the variations of usages of electrical services in a building. It also includes a user interface where the user may provide his plans for the future. The proposed approach is restricted to the prediction of appliance usage, based only on the appliance consumption data and the time of the event.

2.1.2 Home Automation System

A home automation system basically consists of household appliances linked via a communication network allowing interactions for control purposes [Palensky 1997]. Thanks to this network, a load management mechanism can be carried out. It is called for example *distributed control* [Wacks 1993] or energy management system (EMS).

Building automation is traditionally used to increase comfort, to enable remote access to buildings and to increase the efficiency of buildings. These systems may aim at applying the best energy assignment plan and a good compromise between energy production and energy consumption, determined by the local EMS. Figure 2.2 also describes the principle of a home automation system in homes and offices.

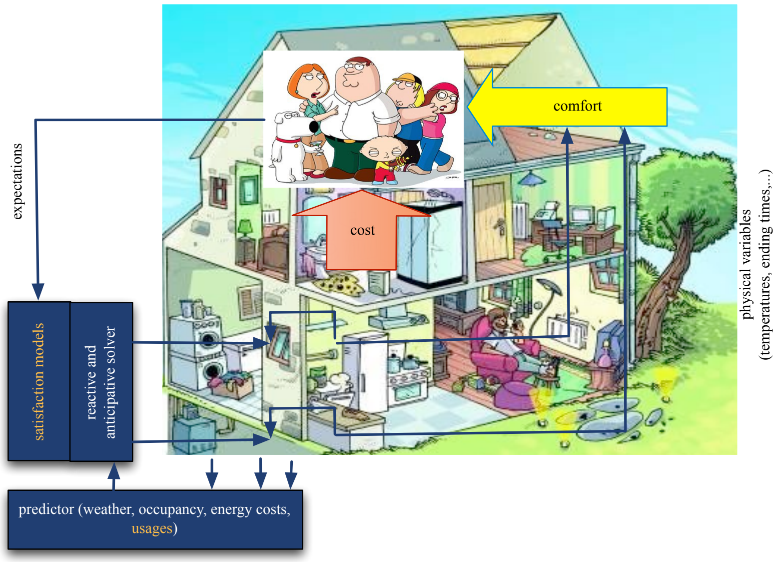


FIGURE 2.2 – Home automation principle

Home automation systems with appliances aiming at providing comfort to inhabitants could base their economic model on services [Humphries 1997]. The modification of consumer demand for energy through various methods such as financial incentives is called demand side management (DSM). The services can be decomposed into three kinds : the end-user services which produce directly comfort to inhabitants, the intermediate services which manage energy storage and the support services which produce electrical power to intermediate and end-user services. Generally, when the home automation system is able to modify the behaviour of a service, this service is qualified as modifiable by the system, for example, the modification of the starting time of a cooking service or the interruption of a washing service, etc. A service is qualified as permanent if its energetic consumption/production/storage covers the whole time range of the energy assignment plan, otherwise, the service is named temporary service.

In a home automation system, the user is not supposed to give the system his expectations (requested services). When the user's demand is not known during a given period, the system must take into account this uncertainty by anticipating the energy needed for services. This helps the system to avoid some problems like peak consumption in this period. Therefore, the behavior of the inhabitant has to be in term modeled and integrated into the home automation system [Hawarah 2010].

In order to keep under control the total amount of consumed energy every hour, and then avoid peak consumptions and minimize the energy cost, the home automation system has to schedule as much as possible the energy consumptions in the most appropriate time periods. For example, the washing machine could be planned before or after the oven in a low energy cost period as far as such a plan satisfies the predicted user's request. The efficiency of the anticipated plan is as good as the prediction of the user's request. Indeed if the actual user's behaviour is far from the predicted one, then the reactive layer has to stop an appliance in order to satisfy the availability energy constraint for example, and schedule this appliance later without any energy cost optimization.

2.2 Predicting energy consumption

Short Term Load Forecasting (STLF) has been already applied at the grid level for some time, for example by the French transmission operator RTE (for "Réseau de transport d'électricité") which proposes an on-line daily electricity consumption prediction of all consumer in France [Rte 2014, Rte 2011]. At the individual appliance level, these techniques have yet to be proven. In fact, the problem of individual appliance usage prediction through consumption data is relatively new. Moreover, STLF uses regressive approaches whereas the proposed method is based on classification but the strategies used in the domain of energy load prediction led to the choice of inputs to the predictor.

A study on the approaches used in load prediction is done in [Feinberg 2005]. The approaches range from using methodologies such as similar day, expert knowledge and linear or nonlinear learning algorithms. In that field, a lot of work has been conducted on the implementation of neural networks in the domain of energy load forecasting [Hippert 2001, Bakirtzis 1996, Park 1991, Khotanzad 1998] and SVM models to predict daily load demand for a chosen amount of time [Chen 2001].

To anticipate the energy needed for a service in a home automation system, the system must take into account the uncertainty which can be provided by the user. In this context, a proper prediction of energy demand in housing sector is of much interest.

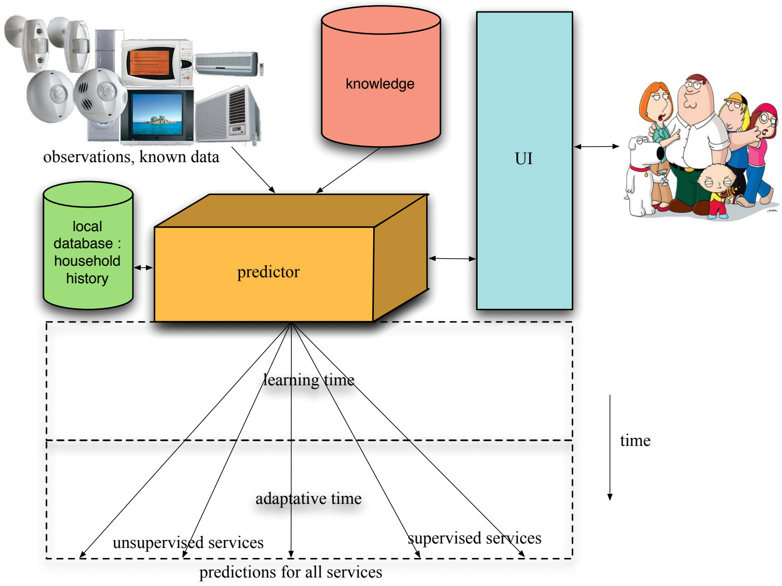


FIGURE 2.3 – Principle of the prediction system

A bottom-up approach can be considered : first, the energy consumption prediction is done for each appliance in a building, then the forecast will be made for the total energy consumed, and finally a prediction can be made regarding the local households supplied by the same energy provider. Indeed, even if it is easier to predict overall consumption, the prediction of the consumption of each appliance is needed for the dynamics of the DSM. It is also important to evaluate how much energy (or if not how much money) can be saved thanks to request to customers like unbalancing requests or energy price variations, these variations depending on particular appliances. The energy savings depend on appliances : some can be unbalanced, some can be postponed and some cannot be changed.

The overall operation of the prediction principle in a Ems is described in Figure 2.3. The residential Ems also includes a user interface where the user may provide his plans for the future. Despite that possibility, in order to limit the work and to stay in restrictive conditions, our proposed approach is limited to the prediction of appliance usage using only appliance consumption data (historical values) and time of the event.

Due to the randomness associated with the use of appliances, regressive model has not been proven very useful in appliance usage prediction. To get over this problem, the solution that we have chosen is to divide the day into 24 hour samples and to predict if the appliance is consuming or not in the time slots. This sampling of the continuous time space in 24 discrete samples makes the prediction more realistic. The details of the proposed method is provided in Chapter 8.

2.3 Energy Management from inside the buildings

The notion of building energy management and control system consists of a set of appliances fitted with micro-controllers able to communicate via standard protocols [Stum 1997]. Various authors have studied control systems dedicated to homes for tracking purpose. For example, [Zhou 2005] and [House 1995] have proposed optimal control strategies for Home Ventilation and Air Conditioning system (Hvac) taking into account the natural thermal storage capacity of buildings that shift the Hvac consumption from peak-period to off-peak period. [Zhou 2005] has shown that this control strategy can save up to 10 % of the electricity cost of a building. However, these approaches do not take into account the energy resource constraints, which generally depend on the autonomy needs of off-grid systems [Muselli 2000] or on the total power production limits of the suppliers in grid connected systems. [Pedersen 2008] have proposed a temperature tracking control using Dynamic Matrix Control (Dmc), a variant of Model Predictive Control (Mpc). But, optimal tracking does not maximize energy usage efficiency.

Some authors have considered in particular the management of local production means and storage systems [Henze 2003, Eynard 2010]. A lot of optimization algorithms have been tested for energy load management [Long 2005, Ha 2006a] including dynamic programming approach [Riffoneau 2010], real-time simulation [Missaoui 2010, Rigo-Mariani 2014] and multi-agent approach [Negenborn 2007]. The general approach of the energy management in living places yields new issues :

Solving energy management problems where uncertainties are predominant - A possible solution (chosen in this work) is a three layer architecture [Ha 2008] which is both able to satisfy the maximum available electrical power constraint and to maximize user satisfaction criteria, through a reactive layer.

Solving large dimension optimization problems - A way of tackling the optimization problem is to use a mixed integer linear programming approach that can manage thousands of binary and continuous variables [Ha 2009].

Solving singular problems - Multi-agent approaches have been used to manage services that can only be modeled by non-linear equations [Abras 2006].

Generating dynamically the energy management problems to solve - Every living place is unique and evolving. Dynamic optimization problem generation has been studied in [Warkozek 2009]. Software architecture and solving process have been depicted in [Ploix 2010]. Figure 2.4 illustrates the proposed solution.

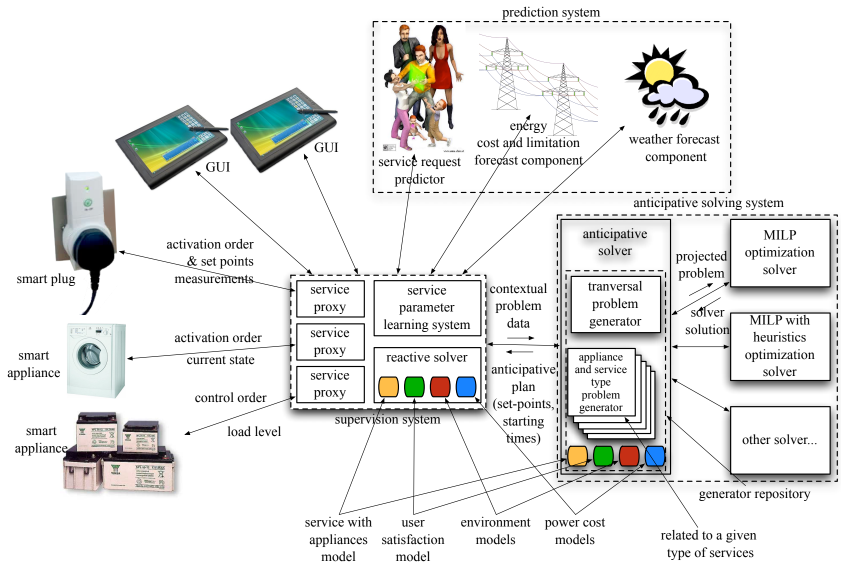


FIGURE 2.4 – Architecture of a power manager

Technical aspects related to communication means within the smart house needs still some standardization. Some details on that subject can be found in [Project 2009]. Hardware and software architectures to predict and optimally manage energy are also the subject of patents [Miller 2006].

Another interesting aspect of the buildings EMS is the real-life experiments. These experiment are the key to making progress in the validation of algorithm and method as we are presenting in this work. As a representative example, the *PlaceLab* is a real home where the routine activities and interactions of everyday home life can be observed, recorded for later analysis, and experimentally manipulated. Volunteer research participants individually live in the *PlaceLab* for days or weeks, treating it as a temporary home. Meanwhile, a detailed description of their activities is recorded by sensing devices integrated into the fabric of the architecture [Intille 2005, Intille 2006b].

Three key challenges must be overcome from such work :

1. Need for comprehensive sensing ;
2. Need for labelled training datasets ;
3. Need for complex, naturalistic environments, to evaluate how typically users will react to a prototype technology in a representative setting

Some responses to these challenges are already being developed, for example, through technologies and design strategies that use context-aware sensing to empower people with information that helps them to make decisions, but without controlling their environment [Intille 2006a].

There are technical and human-computer interface advantages of creating systems that attempt to empower users with information at “teachable moments” rather than automating much decision-making using “smart” control.

Uniqueness of housing systems involves the necessity to develop new tools and algorithms for globally optimized power management of the home appliances, able to anticipate specific situations but also able to take into account the actual housing system state and the occupant expectations. This global control approach leads to the concept of energy smart home, which is more ambitious, than home automation. It should help to keep the balance between consumption and electricity production on the home scale but also at building, neighborhoods and grid scales. Smart homes should be able to take into account external signals, like energy prices or unbalancing orders, and to modify the home appliance behaviors to compromise between occupants’ expectations and external actor wishes.

As an example of work combining hardware and software, three tools for acquiring data about people, their behavior, and their use of technology in natural settings are proposed in [Intille 2003] to make a first step toward a fully monitored smart home :

1. A context aware experience sampling tool, which offers a variety of options for acquiring self-report data from users or subjects in experiments (software)
2. An ubiquitous sensing system that detects environmental changes. It can collect data via measurement of objects in the environment and can complement the self-report data collected by the context-aware experience sampling device (software + hardware)
3. An image-based experience sampling system. this tools combines scene-based sensing and sampling techniques

Moreover, uniqueness also requires cheap installation and maintenance costs because economy of scale is not possible. It means that the new tools and algorithms will have to be easy to install thanks to auto-discovering and auto-learning capabilities, easy to reconfigure and easy to repair. These issues involve sensing capabilities and intuitive human machine interfaces. This field represent a huge opportunity of development for the open-source community, from the hardware, and software point of view (example Zigbee [Baronti 2007]).

A system for recognizing activities in the home setting using a set of small and simple state-change sensors has already been developed [Tapia 2004]. The sensors can be quickly and ubiquitously installed in home environments. This system presents an alternative to sensors that are sometimes perceived as invasive, such as cameras and microphones. In the work the prediction system is based on the energy consumption of the appliances, no other sensors are placed in the house.

Anticipating problematic situations requires prediction capabilities. Weather forecasts have to be fit to real housing environments, taking into account building aspects and masks involving shadings. Occupant behavior should also be predicted in order to avoid interrogating inhabitants about their intended activities. But predicting the use of an oven is a quite difficult problem with no apparent regularity. New prediction algorithms have to be developed where data comes both from history and from the expression of intentions by occupants using proper Human-Machine interfaces. The data privacy issue is in contention for some time now. The unregulated presence of sensors is also a matter of concern. Any research in this direction should have minimal intrusion of privacy and be properly regulated.

Summary

In this chapter, the issues related to the energy management of residence are discussed. The energy management for loads can be both from the gird level or the residence level, intricacies of both the scenarios are discussed. Residential energy management have many aspects to it, starting from anticipation of the behavior either by intrusive or non-intrusive means, followed by large scale control and optimization strategies. The energy management system will be significantly aided by an appliance level prediction system. Various aspects of a prediction system is discussed in the home environment. Another aspect of this kind of system is just the use of appliance consumption data and not using any other sensors, this can significantly contribute in the reduction of user privacy concerns. In the part IV a novel application prediction strategy is implemented and the intricacies of the system addressed.

Deuxième partie

Qualitative Energy Data Analysis

Chapitre 3

Database and Energy Consumption Analysis

3.1 Data-set **25**

3.2 Visualization Techniques **26**

3.2.1 Context 26

3.2.2 Pixel Based Visualization 27

3.2.2.1 Choosing the colors of the pixels 28

3.2.2.2 Arrangement of the pixels 29

3.2.2.3 Windowing the data 29

3.2.2.4 Chart size and shape 30

3.2.2.5 Granularities 31

3.1 Data-set

The data-set used in this work is a sub-set from the results of an European project called *Residential Monitoring to Decrease Energy Use and Carbon Emissions in Europe* (Remodece). This is a database of residential consumptions, from Western, Central and Eastern European Countries, as well as more recent European Countries (like Bulgaria and Romania).

The part of the dataset used is a sub-set called Irise. It is a part of Remodece, dealing only with houses in France. The database consists of energy consumption monitored in 100 households during a whole year. For each house, a data-set consists of the recordings of aggregated power for almost all electric appliances in the house at a sampling time of 10 minutes over a year. In addition the data-set contains weather information at a sampling time of 60 minutes (temperature, humidity etc.), the number of residents, the area and the location of the house.

Fig. 3.1 shows a comparison of the 100 houses grouped by average energy consumption over a year, showing that much more houses are low consuming (<100 Wh) than medium consuming (100-200 Wh) and even less houses are high consuming (>200 Wh).

In Figure 3.2 a comparison of the average energy consumed by deferrable appliances to the total energy consumer is shown for the 100 houses of the database Irise. It can be observed that deferrable appliances also account for a significant amount of consumption in the houses.

From the Figure 3.3, it can be observed that even at a low sampling rate of 10 minutes, appliances have various energy consumption levels and frequency of usages. The most frequently used appliance for this house is the Microwave Oven but it consumes less energy during one

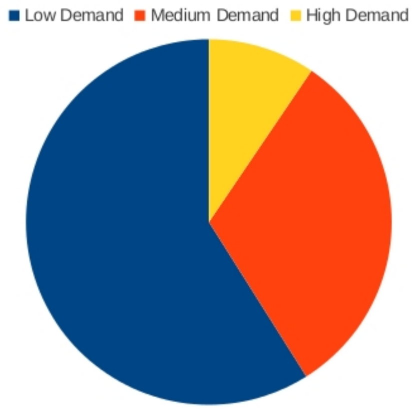


FIGURE 3.1 – Comparison of the 100 houses of the IRISE database by average load consumption over one year of data.

hour compared to the water heater which is less frequently used but consumes significantly high energy. The washing machine varies both in frequency of usage and the energy consumption. The snapshot at the raw data in .csv file format is attached for a particular house is shown in the Appendix A.1.

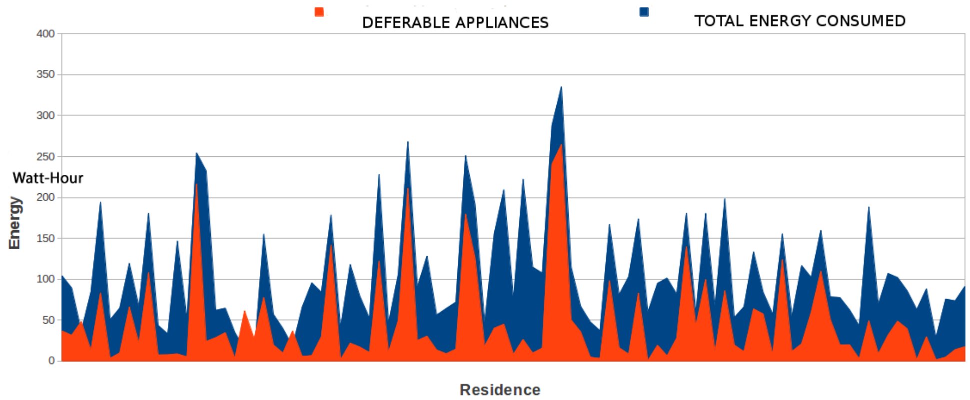


FIGURE 3.2 – Repartition of the consumption of the 100 houses of the IRISE database by average load consumption over one year between all loads and only deferrable loads.

3.2 Visualization Techniques

3.2.1 Context

Advanced data visualization techniques have evolved over the years and represent an important aspect in data interpretation. After all, it is said a figure represents a thousand words ! With the increasing number of smart meters, the quantity of data has also increased exponentially. A new field of research is being explored at that point and is called *Big Data*.

When addressing the smart buildings problematic, it is not sufficient anymore to visualize hourly consumption plot : more advanced visualization techniques are available and better suited.

Visualization techniques can be categorized as static or dynamic, global or local, two di-

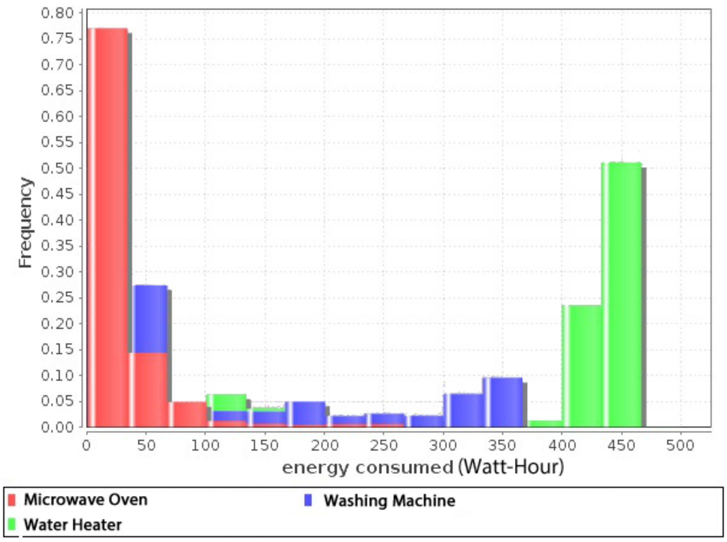


FIGURE 3.3 – Frequency of usage of different appliances within a house for different energy levels at one hour.

mensional or three dimensional, etc. The energy consumption is essentially a uni-variable time series and there are many existing visualization techniques for such temporal data. The task is to identify techniques which will be beneficial from the energy management point of view and also easily interpreted by the user.

To visualize the data dynamically and locally in two dimensions, a cluster and a calendar based visualization could be used. These graphics help to visualize the energy consumption at different level of aggregation. This way of visualization works well for interpretation but it is difficult to compare between houses. In order to work around this problem, we used a pixel based method for energy data visualization.

3.2.2 Pixel Based Visualization

The pixel based visualization method allows representing a large amount of information in a graphic of limited size. For this purpose, each data is represented by a pixel whose color corresponds to the value of this data. Over a two-dimensional spatial organization, a pixel based representation is required in order to view and differentiate the data from each other. For this purpose various forms can be chosen as shown in this section.

Such a representation method has several advantages. It allows the representation of a very large number of data in one graph (up to the millions of values). This is ideal in the study case of energy consumption monitoring in buildings, or group of buildings : as they are explored on smart buildings and smart cities.

Due to the large amount of data shown, these visualization methods allow also the detection of recursive schemes. In the observation of a phenomenon relative to one dimension (e.g. time), the recursion is when a periodicity or a trend can be observed between the phenomenon and one or more scales of the considered dimension. E.g. for energy consumption, one can observe an increase in consumption between 18 h and 22 h for a majority of households, there exists a recursive scheme between power consumption and the time scale of hours.

To achieve a pixel-based visualization, several general parameters are to be determined : the color scale (for color mapping), the arrangement of pixels (for pixel arrangement), the chart size (for the ordering of the dimensions) and finally the shape of the sub-windows.

To apply this method into a meaningful system, specific parameters have to be considered. The parameters consist of the color model, the arrangement of pixels, the shape of the window and the granularity. Indeed, if the representation area is poorly designed, then it is certain that the user cannot observe any recursive scheme and therefore be lead to wrong conclusions. Each representation must be made case by case depending on the nature of the observed phenomenon. There is no generic method for all applications. This method therefore requires careful assessment of all the relevant parameters.

The energy consumption data will be represented here only as a function of time. Consumption is measured with an interval of 60 minutes. In the Figure 3.6, the representation of data over a year for one house is shown at different granularities.

In the Figure 3.4, two days from the same house are compared to observe the various patterns of consumption.

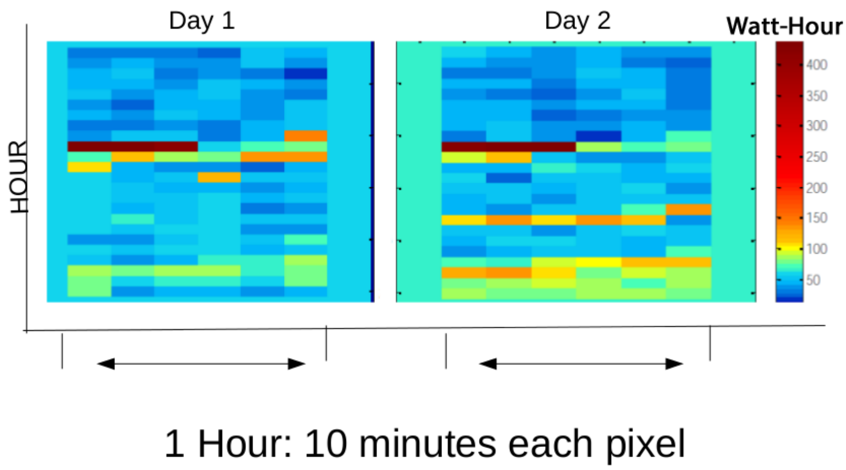


FIGURE 3.4 – Day to day comparison using pixel based visualization.

3.2.2.1 Choosing the colors of the pixels

As explained above, the data value is represented by the color of the pixel. Subsequently, the task is to determine the most appropriate scale for an easy reading of the chart. The task is not only to select color but the shades that are most visible to the human eye. It is therefore a question of principle “just noticeable differences”.

The goal is to have the largest number of colors well separated and as distinguishable as possible. To determine the easiest and most common color, the task is to vary different characteristics that separate colors : intensity, brightness and saturation. This is called the HSV (Hue – Saturation - Value) system. It is one of the most used in the case of graphical representation. Experimentally, the most effective set of colors is the one that uses all the colors of the rainbow scale. This is indeed the most intuitive to the human eye in term of color range.

Once the color range (i.e. the gradient) is chosen, the task is to determine a scale with a specific number of colors to represent the data. Each color should be chosen to correspond to a range of values.

The chosen color scale is reported in the Figure 3.5 and described as follows :

- Basic scale using the gradient colors of the rainbow but can be used in several different scales and should be implemented in a way that the user has a choice.
- To schedule representation, an option has to be provided that allows the color of each pixel to be determined in part by the average value of the subdivision (together with an adjustable rate of opacity).

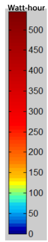


FIGURE 3.5 – Colour Model for Visualization.

3.2.2.2 Arrangement of the pixels

Since the requirement is to visualize huge amount of data, the arrangement of pixels is essential to unravel recursive schemes. Without a suitable arrangement, the graph would be a multitude of patterns of seeming colors that are arranged randomly and therefore unusable.

The arrangement depends on the nature of the data and thereby is domain specific. In addition, the interval between each measurement is also taken into account (for example energy measurement every 60 minutes). There is no general mathematical rule cutting across domains. It is nevertheless necessary to propose a mathematical relationship for each case in order to automatically position the pixels in the graph.

In the case of this work, we consider only data having a natural order of arrangement. Indeed, the considered data correspond to temporal power readings. And two adjacent data over time also will be placed in the adjacent pixels.

3.2.2.3 Windowing the data

Instead of representing all data in a single rectangle or in a single block, It is better suited to subdivide the area of representation. This temporal subdivision can highlight recursive schemes to smaller scales while allowing a global view of all values over the period. The advantage of using sub-sampling is the possibility of visualization in-terms of daily, weekly, monthly and seasonal patterns over a year. In the Figure 3.6 these patterns are highlighted from the model presented in [Lammarsch 2009].

In addition, the use of sub-windows can allow the representation of additional data, such as average value over the interval represented by this pane. Each subdivision of the representation

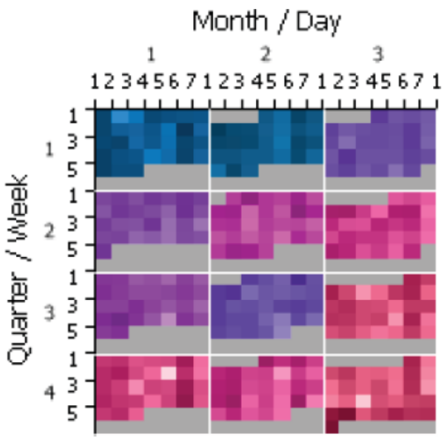


FIGURE 3.6 – 1-year representation model of the data with four different granularities.

allows performing calculations for smaller scales. This subdivision can then impact the color of the pixels.

The use of opacity is also possible [Lammarsch 2009]. The color of each pixel is determined both by the average color of the subdivision and share the precise value of the data corresponding to that pixel. By varying the relative opacity of each of these two sources, the user can have a broader view and more precise data.

In fig. 3.7 the recursive nature of energy data is observed by fixing the window size to twenty eight days and then compared to a case where a random window size has been used.

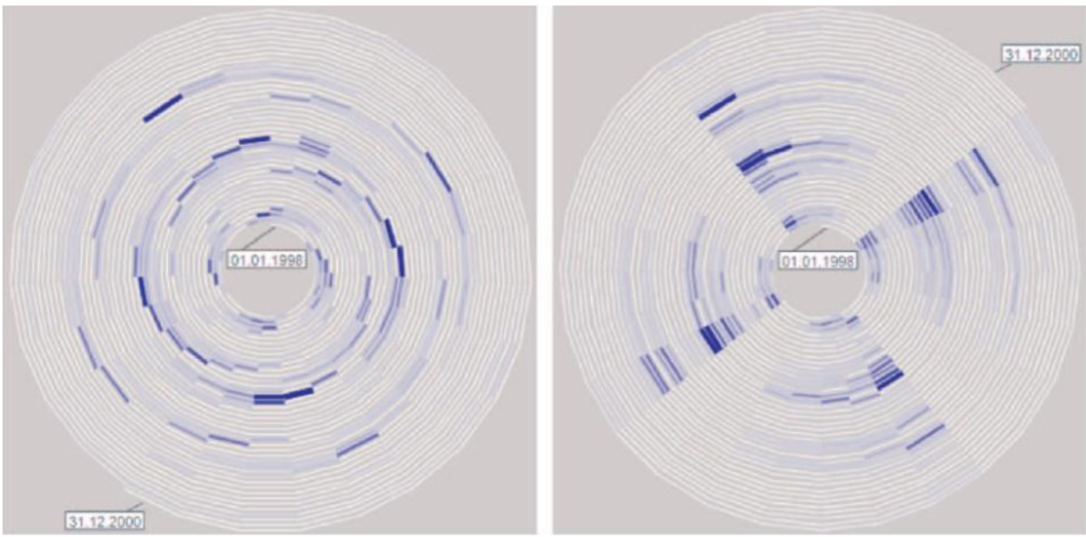


FIGURE 3.7 – 28 days window size (left) compared to random window size (right).

3.2.2.4 Chart size and shape

When it comes to creating a graph, the question of axes dimensions arises. A graph can be interpreted as if a map is shown (two dimensional) or in the space (three dimensional). The objective here is to be able to stay in the plane thereby in two dimension. Several solutions are possible :

– If two lines are sufficient for arranging all data, then a rectangular shape will be ideal (note : considering the granularity).
– If more than two axes are needed to match all data then a polygon representation could be used where each axes represents a edge of a polygon. This representation allows highlighting particular cyclic repetitive patterns.

3.2.2.5 Granularities

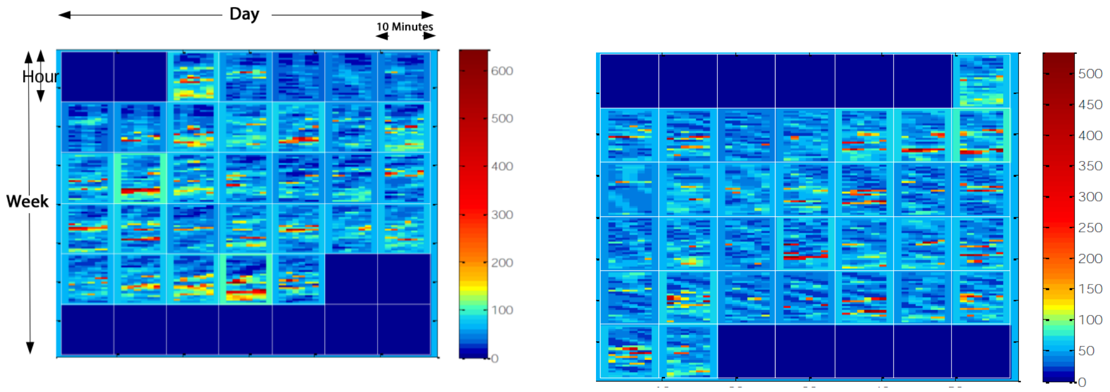
Usually the data are represented in terms of a single parameter (e.g. time). Nevertheless, it is possible to subdivide this parameter into smaller parts : this is known as granularity.

Granularities can be different for each axes of the graph. Furthermore the same axis can represent two different granularities. For example, if we consider the representation of the time : a first axis could represent months while a second one may represent days. Two different granularities of the same parameter are used to arrange the represented data. Each axis can both determine the month and day of the week in question.

The number of granularity to use will determines the number of axis required. In the Figure 3.8, data are represented by a function of time. Multiple granularities have been used :

On the horizontal axis : the day and 10 minutes.

On the vertical axis : the week and hour.



(a) Month 1. (b) Month 2.

FIGURE 3.8 – Monthly energy consumption representation with hourly and daily granularities.

The graph will be observed in two axes, each of them having up to two granularities. The user must choose the configuration they want, i.e. choose the total measurement time (one day to one year) and also chose the number of granularities that has to be taken into account in the representation. Different granularities are available : 10 min, one day, one week, one month and finally one season. The graph will then show several areas depending on the chosen granularity : respectively minutes, hours, days, weeks or months.

Three granularities are proposed in figures here and are defined as follows :

1. Per day visualization, illustration in the Figure 3.4.

2. Months visualization, illustration in the Figure 3.8.
3. Visualization by year, illustration in the Figure 3.9.

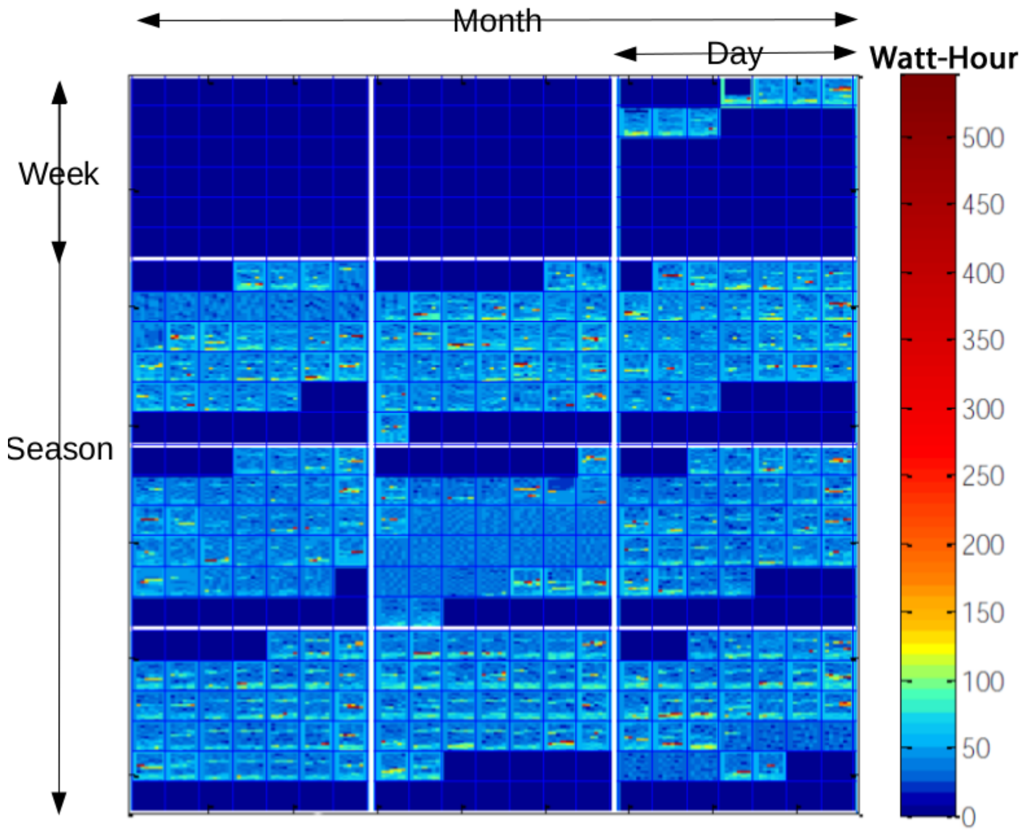


FIGURE 3.9 – Yearly Energy consumption with hourly, daily, weekly and monthly granularities.

In the figures 3.4, 3.8 and 3.9 the data can be observed at different scale and granularity. This will allow the user to identify his peak consumption period is various time scales (daily, weekly, monthly and so on). A color model based on opacity can add information about the average consumption over a period (e.g. monthly). Note that this way of presenting data is completely compatible for use on a touch-screen, including zoom from one granularity to the other.

Finally, the consumption information may be loads in a building, a residence, a group or at the distribution grid level or a cluster of houses. The level of the application and the intended beneficiary will determine the final implementation of the data visualization system.

Conclusion

The dataset used in the thesis is discussed in details with qualitative analysis of the consumption patterns. The dataset consists of 100 houses consumption data at sampling rate of 10-60 minutes at both appliance and meter level. Weather and general consumer information is also available. Various statistical analyses on the dataset is observed to understand the dataset and also the importance of deferrable appliances. Big data visualization is an important aspect of modern day research. In this work, a relevant pixel based visualization method is discussed and implemented for energy data visualization. The data is visualized at various granularities to observe different temporal pattern within the data. This can be a tool for both the grid manager and the consumer with different requirements. The application could be used at different levels starting bottom-up from the appliances to the grid energy consumption.

Chapitre 4

Clustering and Flexibility

4.1 Categories of households **35**
4.1.1 Clustering using K-means 35
4.1.2 Principal Component Analysis 36
4.1.3 Features and clusters 36
4.2 Categories of appliances **37**
4.3 Load flexibility for the grids **39**
4.3.1 Flexibility of specific loads 40
4.3.2 Assessing flexibility of a group of houses 41

4.1 Categories of households

Clustering is the unsupervised learning task of grouping a set of objects into different categories. The objects of the same group are more similar to each other than the object in the other groups. In the following section, the clustering principle and data dimension reduction algorithm are discussed.

4.1.1 Clustering using K-means

K-means clustering [Hartigan 1979] is used to partition a set of observations into a set of clusters in which each object belongs to the cluster with the nearest mean. It effectively partitions the data into "Voronoi" cells which is a way of dividing space into a number of regions.

In the classical method the number of partitions K had to be provided but many variants have evolved to overcome this restriction. In our work, an X-means algorithm is used. This algorithm extends the K-means principal by an Improve-Structure part. In this part of the algorithm the centers are attempted to be split in its region. The input vector is normalized using the min-max algorithm.

Given a set of observations $(x_1,\ x_2,\ \ldots,\ x_n)$, where each observation is a d-dimensional real vector, K-means clustering aims to partition the n observations into $K \leq n$ sets $S = \{S_1, S_2, \ldots, S_k\}$ so as to minimize the within-cluster sum of squares. In other words, its objective is to find :

$$\arg\min_{s} \sum_{i=1}^{K} \sum_{x_j \in S_i} ||x_j - \mu_i||^2 \tag{4.1}$$

where μ_i is the mean of the points contained in the set S_i.

4.1.2 Principal Component Analysis

Principal component analysis (PCA) [Jolliffe 2005] also known as orthogonal linear transformation involves a statistical task that transforms a number of correlated variables into a smaller number of uncorrelated variables called principal components. Mathematically, PCA solves eigenvalues and eigenvectors of a square symmetric matrix with sums of squares and cross product. The eigenvector associated with the largest eigenvalue is the direction of the first principal component and henceforth.

Physically, the first principal component represent the largest variability in the data and each following component is used for the remaining variabilities. PCA is used in many fields for different purposes. In this work it is used to reduce the dimension of the data for easier visualization and understanding without significant loss of variability. It is also used in the chapter 8 for the dimensional reduction of the input vector.

4.1.3 Features and clusters

To make qualitative analysis of the data-set, a clustering of the 100 houses of the IRISE database has been done. The results obtained using a X-means clustering analysis resulted in four categories of houses. The features taken into account for the analysis were the mean and standard deviation of the energy consumed over one hour, the number of residents, the area of the house, the number of deferrable appliances in the house, the mean of deferrable appliances over a year (refer to the Table 4.1) and the number of usage of the appliances for different energy levels (refer to the Figure 4.1).

TABLE 4.1 – Features for residential data analysis.

Cluster	Mean of energy	Standard deviation	Number of people	Area of the house	Number of defer. load	Mean of defer. load
1	234	771	3.8	119	6.7	190
2	105	180	4	113	4.1	35
3	214	319	3.4	109	4.3	111
4	56	116	3.1	96	3.1	17

The cluster centers for the four categories are proposed in the table 4.1 along with some features, in the Figures 4.1 and 4.2 for visualization.

It can be observed in the Table 4.1 that the increase in average hourly energy consumption is not only proportional to the number of deferrable appliances present in the house but also depends on the number of people present in the house.

The Figure 4.1 proposes the frequency of usage of the loads in each house in relation to energy levels, from 0 to more than 400 kW.h. All algorithms and methods proposed in this work are tested on houses representative of one of these defined categories.

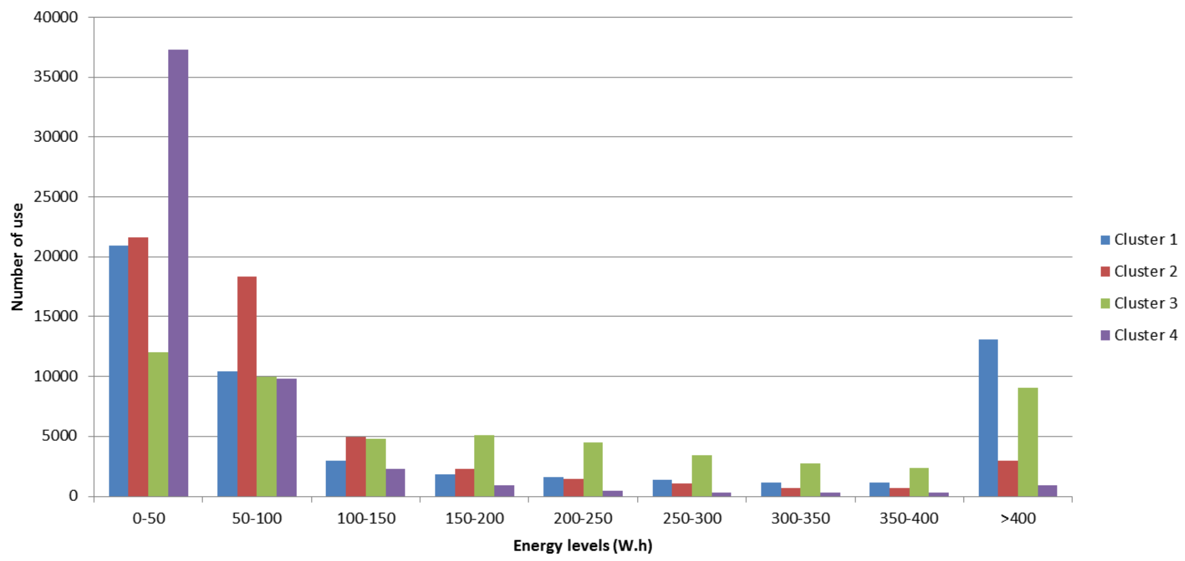


FIGURE 4.1 – Number of usages of the appliances for different energy levels.

In Figure 4.2 the 100 houses are projected using the features shown in Table 4.1 and in Figure 4.1 (features values for the four cluster centers). The colors and the shapes represent the clusters and the size (surface area) of any point is representative of the hourly mean energy consumed over a year.

From the projection of the houses into two principal axes using principal component analysis as shown in Figure 4.2, the houses are grouped in different categories based on their number and use of appliances. This point is discussed in the Chapter 7. The parallel coordinate plot for the houses in the dataset and the cluster centre is observed in annexture A.2.

4.2 Categories of appliances

The number of consumer appliances in buildings is increasing continually ; therefore, it is difficult to review them individually. For that reason, the appliances are categorized as follows :

Continuously working devices - These devices are plugged in all the time during the day. There is a cost associated with these devices if they are turned off (for example the heating devices).

ON/OFF appliances - This category represent most of the appliances in household which have two states (ON and OFF) such as simple water heater and light bulbs.

Multi-state appliances - Appliances that pass through several definite switching states like they could be defined in a finite state machine. Example : washing machines and clothes dryers.

Continuously variable appliances - Devices that have an infinite number of states with a variable power draw. Examples : light dimmers and some power tools.

In the restriction of our work, based on the limitation of current smart meters, the focus is primarily done on high consuming ON/OFF appliances and multi-state appliances. The ap-

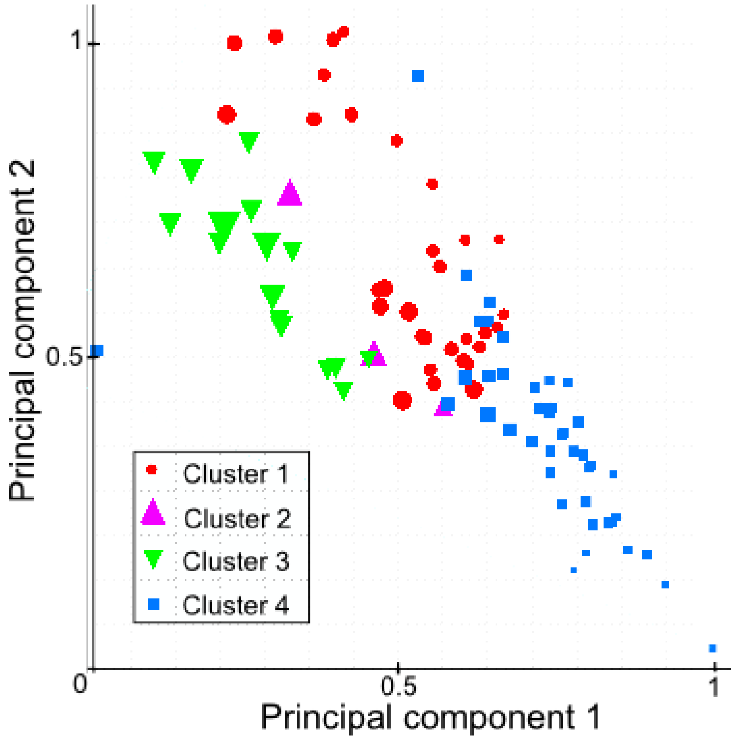


FIGURE 4.2 – 100 houses projected into two principal axis (using PCA) and divided into 4 clusters

pliances classification allows a better understanding of the optimal scheduling problem that will be discussed in the last chapter. For that reason, the appliances are categorized in three main categories, depending on their usability for energy management (or a future *flexibility*) as follows :

1. The plug-in or non-controllable appliances are used for direct consumption and their usage cannot be controlled by the grid. This category consists of lighting, TVs, small kitchen and bathroom devices such as microwave, coffee machine, hair dryer and so on.

2. The controllable or manageable loads can be split into interruptible loads, non-interruptible or semi-controllable ones. Hence, their usage can be controlled. Non-interruptible devices can be re-scheduled or shifted in time, but once their work cycle has been started it is more cost effective to complete the usage, an in-between stop and start consuming a higher amount of energy or being simply impossible for the considered device. For non-interruptible devices, it is important to know the consumers preferences in order to design an optimal energy management system which takes into account comfort criteria. This category regroups the major appliances settled by the inhabitants, who can decide to start them immediately or not.

3. The interruptible device has been segmented in multi-modules and permanent switching. A multi-module device means that its working cycle can be seen as a sum of individual sub-working cycles, dissociated over the time horizon. The interruptible devices can be controlled based on constraints which are specific for each concerned device.

4.3 Load flexibility for the grids

Works allowing the grids to become more active and flexible are intended to let the supply of electricity to be more efficient, sustainable, and economical in accordance with security conditions.

The main objective is to find the flexibility of the electrical consumption in a distribution network, for the purposes of increasing or decreasing consumption in relation to a more variable production. The evolution of the production is often related to the increase of renewable energy production that will lead to more intermittent energy production injected in grids, and therefore more difficulties to keep the equality between production and consumption.

This work is mainly focusing on the consumption of one building (in fact going inside it by distinguishing between appliances) but can totally be taken one step beyond by considering the consumption of several buildings, i.e. a district, then a city, etc. That is what will be considered in this section.

The notion of *flexibility* in this context represents the amount of energy that can be removed or added (i.e. increased or decreased) at one point of the time in relation to the forecast of the energy that should be consumed by an aggregation of houses. The aggregation of houses is made on the basis of the clustering shown in section 4.1.3. The objective is to observe the impact on grids by multiplying the amount of energy controlled through the multiplication of the same charge (i.e. the same consumption behavior). This has to be done while keeping the same amount of energy consumption over a given period of time (one day for example). Thef "artificial" variations of the loads (i.e. the differences with a forecast due to an action of control and not a prediction error) have then to be done on smaller time periods than the duration taken as limit for the energy balance. In our case, the time duration for energy balance is of one day, and then the actions on loads have to be in a range of a few minutes to less than 24 hours, including the duration of service of the considered load.

The idea behind these manipulations is more to adapt the energy consumption to energy production than to decrease globally energy consumption. The later observation is related to the former one, but will be addressed in a different manner, effectively through awareness campaign and more probably through prices incentives.

The *flexibility* of the means of consumption, in other words the variation of the energy consumption relatively to its forecast, can be created by a house or a group of houses. The evolution of the consumption compared to a previous forecast (or historical data) can be obtained by load shedding in order to find the minimum consumption or on the contrary by load shifting in order to reach the maximum consumption (while respecting certain constraints).

To start with, one must determine the possibilities of flexibility for one single house, starting at a certain time. This flexibility is achieved by modifying the working cycles of the loads or by rescheduling various categories of loads : movable loads, interruptible loads and modifiable loads. Therefore, the daily *flexibility* in the energy consumption is primarily dependent on the appliances being used during that day, the day before and the next day.

The constraints that can be listed for a possible use of the loads are as follows :

– The total power of the loads does not exceed a fixed limit.

- The energy consumption during one day will remain the same as the energy consumption forecast, regardless the way in which it will be distributed over the day.
- A movable loads cycle cannot be interrupted if it has started.
- The actions on loads are restricted by the inhabitant comfort criteria (are concerned mainly the electric heater and water heater).
- Some appliances cannot start before the end of the cycle of other appliances (for example the clothes drier after the washing machine).
- Some appliances cannot start unless the inhabitants ask them to (for example the washing machine cannot start if there is no clothes inside).
- Some appliances has to end or start before a certain time (for example the washing machine should not run after 10 O'Clock).

All these constraints, and probably many other specifics demands from inhabitants, will have to be taken into account in a global energy management system in future smart buildings, and later in the control and monitoring of smart cities. This kind of energy management system will create a massive amount of private data, the challenge being double : to handle them properly, first, technically and second, in a real secure way.

As a first proof of concept, the objective of these sections is to evaluate the principle of flexibility of various loads in groups of houses through their energy consumption.

4.3.1 Flexibility of specific loads

We consider a scenario where a distribution network operator, through flexibility requests to a specific market (not presently completely functional) wants to act on the energy consumption of a house or of a group of houses, grouped by the similarities in their consumption pattern. In France the regulation is down in NEBEF ([1]) for load shedding. This action is defined as a certain amount of energy to add or subtract (a capacity) during a certain amount of time.

In the work, the time interval is a duration starting at a time of the day where we want to achieve a given scenario, for example, to maximize or minimize the consumption. This interval has a duration of one hour (but could be bigger or smaller) which corresponds to six time slots of ten minutes (10 minutes being the sampling rate of the Irise database).

The situation in the case of the considered loads is simple. They cannot be moved unless their start-up time belongs to the interval defined in the scenario or is predicted to start after (in the case of maximizing the energy consumption). These categories of appliances can be differed or shifted only if the appliances are predicted to be used during the interval of consideration and also that the predicted end of service (taking into account the new starting time) is included in the time period of energy balance.

In the results shown in section 4.3.2, the duration is considered to be of 24 hours. It can be observed that if the appliances are predicted to be used during the end of the day, the flexibility within the day is then reduced. The consumption is moved in totality or by blocks. Taking the inhabitants comfort criteria into consideration, the appliances should only be moved within a restricted period. For example, for the washing machine, this interval has been fixed to 12 hours and for the water heater to two hours.

1. http://www.erdf.fr/ERDF_NEBEF

Another particular appliance is the electrical heating. This load is the largest one during winter days in houses with water heater. It is subject to the restriction, that the heating cannot be interrupted more than 30 minutes. This restriction is an average value based on thermal time constant that could be found in typical rooms in households as they were consider in the IRISE database. The total energy consumed by the heating system is also fixed. The distribution can be rearranged without changing the overall consumption. An additional restriction of maximum shift of 80% consumption is also imposed.

4.3.2 Assessing flexibility of a group of houses

The action of aggregating houses together depending on similarities of consumption patterns is something that has already began to be considered by some industrial actors, like Voltalis [2] in France. These actors will aggregate customers in order to bet blocs of energy that can be shifted in time on capacity markets.

In the graphics shown below, the objective is to assess the flexibility limits (maximum and minimum) that can be predicted for groups of houses, based on the previous knowledge of their consumption pattern (consumption prediction after load identification).

At one point in time, the upper bound of the flexibility is defined when all loads that can be started right away are shifted from their predicted (future) time of use to the current time period. The lower bound of the flexibility is obtained when all loads that can be stopped or postponed are shifted away in the future. This flexibility is assessed under the constraints listed in the previous section which depends on the category of the load, the predicted starting and ending time, etc.

All the houses from the IRISE database are grouped into four categories based on a clustering algorithm (X-means) as discussed previously in section 4.1.3. All houses from one cluster are aggregated together to constitute a group. The appliances considered are the washing machine, the water heater, the electrical heating, the clothes drier and the dish washer. For each group (from group 1 to group 4) three characteristic curves are observed : total consumption, minimum of flexibility and maximum of flexibility.

In the Figure 4.3 the minimum of flexibility of the third group of houses is observed in the smallest cluster described in section 4.1.3. This group contains only three houses that are aggregated together to observe the global possibilities of flexibility. It can be observed by the definition of maximum and minimum flexibility that their value depends on the consumption during that period.

In a similar manner, the Figure 4.4 proposes the maximum of flexibility for the same group (same cluster). The flexibility is reduced during the evening for the considered day due to the general preference of the inhabitants to use the deferrable appliances later during that day. It can be observed for both figure 4.3 and 4.4 that the individual houses have similar consumption and flexibility pattern within the group. It can further be observed that the flexibility is highest during the afternoon.

The minimum and maximum of flexibility proposed by two other groups of houses based on

2. `http://www.voltalis.com`

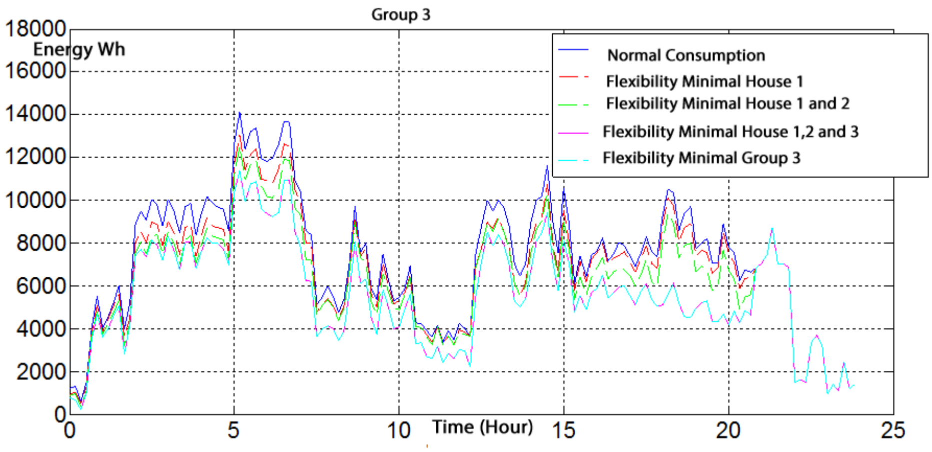


FIGURE 4.3 – Minimum of flexibility of the group 3.

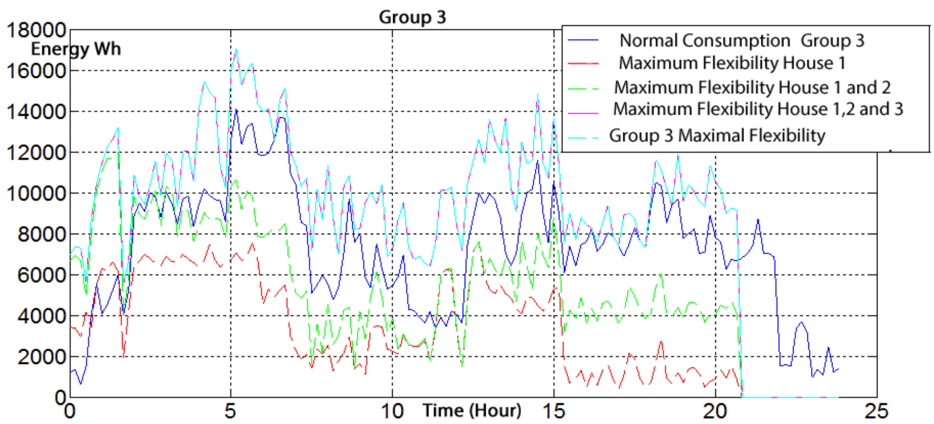


FIGURE 4.4 – Maximum of flexibility of the group 3.

the previous clusters can be seen in the Figure 4.5.

The flexibility is higher for the group 1 than group 3. This is primarily due to the fact that more controllable appliances are present in the group (based on the category of houses with multiple appliances). It can further be observed that the flexibility in the group 1 is highest during the start of the day and reduces as the day progresses which is unlike the case of the group 3. This is primarily due to the difference in nature of the deferrable or controllable appliances present within the groups.

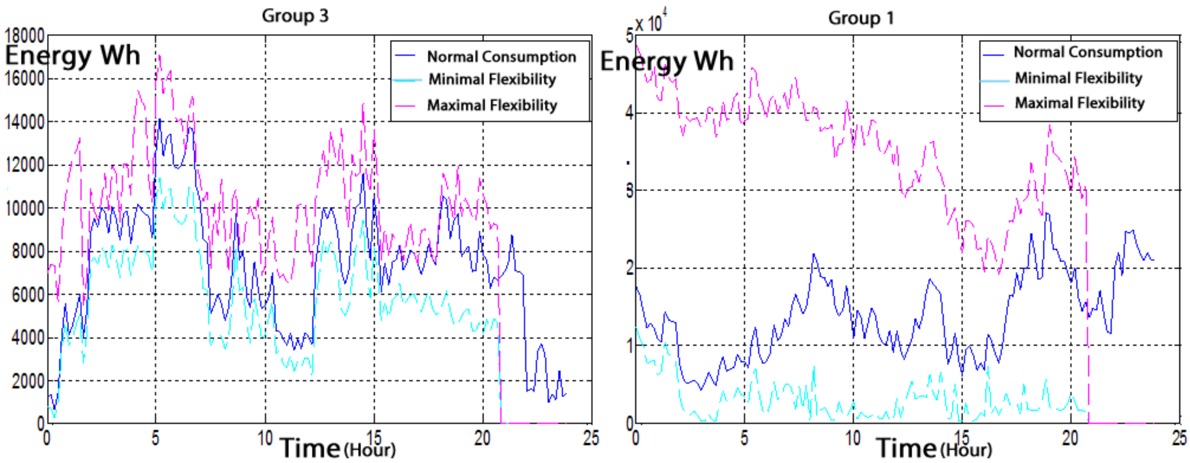


FIGURE 4.5 – Minimum and maximum of flexibility for group 1 (left) and group 2 (right).

A similar graphic is show by comparing the last two groups of houses, group 3 and group 4. It can also be observed from Figure 4.6 that the flexibility reduces as the day progresses. This analysis also provides a bandwidth (upper bound and lower bound) within which an optimal load scheduling algorithm will work. The group 2 has a more stable flexibility throughout the day than group 4, the possible reason is that within group 2 the number of interruptible appliances are more than group 4. The non-interruptible appliances impose restriction which reduces the flexibility as the day progresses.

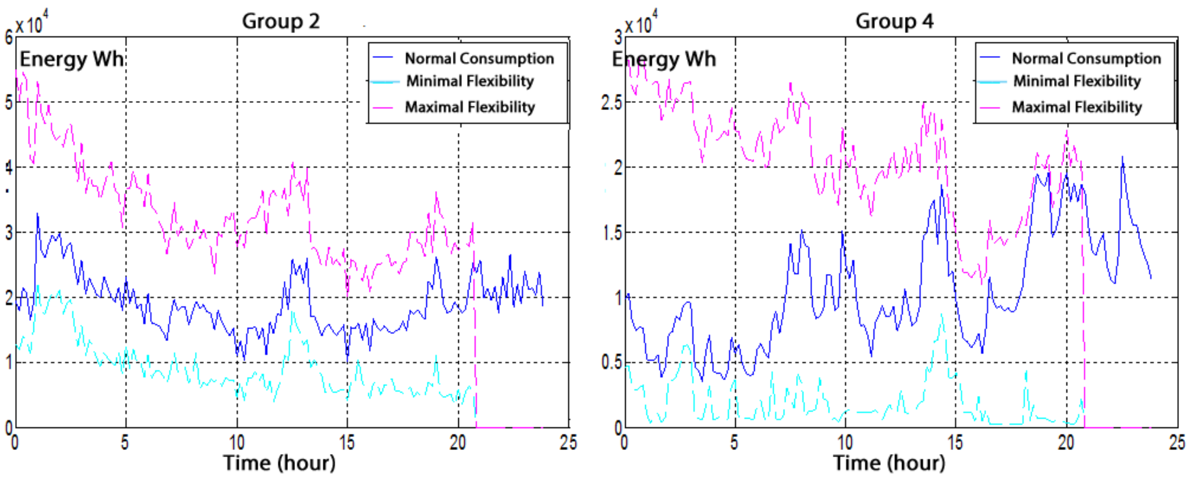


FIGURE 4.6 – Minimum and maximum of flexibility for group 3 (left) and group 4 (right).

The two Figures 4.5 and 4.6 show relatively similar trends in their flexibility, illustrating the fact that the IRISE database is not diversified enough to propose a complete assessment of energy consumption patterns. In fact, even if the houses are in different groups, their number of inhabitant, appliances and load curves are not significantly different. The primary difference is in the nature of the controllable appliances whether they are interruptible or not and the constraints governing them.

As a comment, we should emphasize that these graphics are a preliminary analysis within the database and are being scaled up for further analysis. The appliances constraints related curves could also give an insight into the tradeoff between the user comfort and the possible flexibility.

Conclusion

The IRISE data-set is first introduced and some qualitative results are proposed. The objective is to have an analysis on the nature of the dataset. A clustering technique for the houses is discussed and the issue of the clusters categorization depending on the loads is emphasized. The K-means clustering technique is detailed and a modified form of K-means (X-means) is used for household segmentation. The Principal Component Analysis is briefly presented and used for data dimension reduction and easier visual analysis. All the hundred houses are observed with the surface area of each point representing the mean hourly energy consumed over a year. The different categories of appliances are discussed and the different types of controllable loads are detailed. Finally, the concept of the flexibility of the energy consumption of buildings is discussed, followed by an illustration of its upper and lower bounds on groups composed by the aggregation of all houses from each clusters. It is observed from the results that the number of controllable appliance play an important role in the possible flexibility in load consumption. The load consumption is more flexible with the presence interruptible appliances such as Heating or Water Heater than with non-interruptible controllable appliances. This is a preliminary work in this regard, it needs to be further incorporated with user defined constraints and observed over the whole year rather than the one particular day. The load flexibility can be observed also in a pixel based visualization to observe the high and low flexibility period.

Troisième partie

Non-Intrusive Load Monitoring

Chapitre 5

Methodology of Temporal Classification

5.1 Temporal sliding window **48**
5.2 Meta-features **50**
5.2.1 Principle 50
5.2.2 Proposed Meta-features 51
5.2.2.1 Hour of the day 52
5.2.2.2 Distances from the current event to the local maximum and local minimum 52
5.2.2.3 Energy variation between time steps 54
5.2.2.4 Gradient and Laplacian of energy consumption within the window 54
5.2.2.5 Mean and standard deviation 54

Introduction

The methodologies to solve the Nialm problem encompasses a mixture of domains. A majority of the earliest research focused on this problem from a signal processing perspective. The focus was on identifying different appliance signatures which distinguishes one appliance from another by analysing with mathematical tools (for example wavelet transformation) [Bier 2013].

Subsequent research also considered the problem as a blind source separation task and proposed relevant techniques in that direction [Kolter 2010]. Our work considers this problem as a temporal classification problem as summarized in the Figure 5.1. The advances in temporal classification have been proven useful in other research domains. As an example, temporal classification is used to predict protein secondary structure from the protein's sequence of amino acid residues and for text mining [Eddy 1998].

The present work formalizes a generic appliance identification technique based on a *multi-label learning process* using a temporal windowing approach where the only input after the training phase is the *time stamped energy readings from the power meter*. The implementation of the load identification technique is described below, while each step is discussed in the following sections.

1. The energy readings are extracted from the Irise dataset at the sampling rate of 10 to 60 minutes.
2. Sub-sequences are generated from this dataset using temporal sliding windows (refer to Section 5.1) with a window size of 10 units.
3. Meta-features are computed for each sub-sequence (refer to Section 5.2).

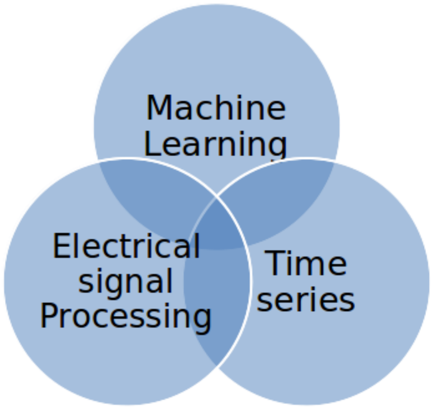


FIGURE 5.1 – Methodology for temporal classification, domains of interest

4. The features thus generated are processed as input attributes and the high energy appliances as output classes for the multi-label classifier (refer to the Chapter 6).

Note, the model is trained using 10 % of the dataset and evaluated on the remaining dataset. This restriction is imposed to have a minimum training period and observe the impact on unknown test instances.

Considering a 10 minutes sampling rate, the method can be directly implemented using current smart meters technologies without any other energy boxes. It reduces the privacy concerns as the daily user activity is less detectable. Indeed, only the high consuming events leave a footprint (even if the event is very short in time, a high consumption will be detected in the energy trace).

The drawback of this approach is that short term and long term events with low power consumption remain undetected.

5.1 Temporal sliding window

We introduce a few key-terms to facilitate the understanding of this work. Temporal data mining encompasses time series analysis on the form, type and scope of the data. The temporal data can be represented by time series or events and can be processed with tools such as classification among others. Some definitions are summarized below :

Time Series : An ordered set of n real-valued variables $T = t_1, \ldots, t_n$.

Sub-sequence : For a given time series T of length n, a sub-sequence C_k of T is a sampling of length $w \leq n$ of contiguous positions from T, that is $C_k = t_k, \ldots, t_{k+w-1}$ for $1 \leq k \leq n - w + 1$.

Time Sliding Window : Given a time series T of length n and a sub-sequence S_t of length w of that time series, a matrix M of all possible sub-sequences can be built by “sliding windows” across T and placing sub-sequence C_k in the k^{th} row of M. The size of the matrix M is $(n - w + 1) \times w$.

In the field of load identification in households, the input (energy) is a time series with an ordered set of real-valued variables whereas the output (predicted classes) is an ordered set of events (appliance states). In this work, time series sub-sequences are generated from the energy reading, and then meta-features are extracted from the sub-sequences to identify the appliances states. The classifier system for load identification is based on temporal classification using standard propositional machine learning algorithms.

The initial step is to populate the sliding window with sufficient historical data that aims at creating a single test instance to start the closed loop classifying process for the future time steps (priming). The subsets of the original time series are then shifted in time creating thereby the sub-sequences and preserving time dependency among sub-sequences. Instances containing these sub-sequences are finally presented as standard propositional instances to the classification algorithm. This process is illustrated in the Figure 5.2. To summarize, the input of the classification algorithm is the subsequences of the total energy consumption and the output is the appliances states (in the Figure 5.2, the appliances are the washing machine and the clothes drier).

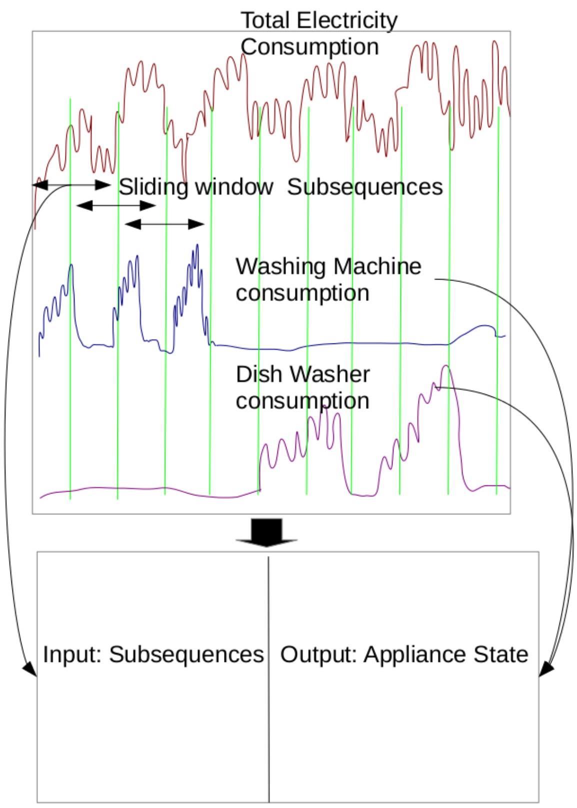


FIGURE 5.2 – Data representation in a standard propositional learning format

As the variables in each sub-sequence are considered as independent the time dependency is lost. The temporal sliding window principle injects back this dependency among the sub-sequences. The time dependency of the usage of the appliances is taken into account but in a different way than event detection at low sampling.

Once the classifier produces a prediction for the next time step, this classified value moves into the sliding time window as the most recent value of the target and the oldest value in the window falls out. Another test instance is then created from the history window and the next time step is classified. This is also known in the literature as “closed loop forecasting”. Once the classifier is trained, it only needs to be “primed” each time for classifying future instances.

The size of the sliding window is fixed experimentally to 10 units, each unit being of 10, 30, or 60 minutes, depending on the sampling rate. Increasing the window size can increase the complexity of the algorithm (not always with a visible change in performance) and decreasing the window size can lower the performances. After experiments, the window size can be reduced up to five units for a lower sampling rate (30 to 60 minutes) without a significant drop in performance. The minimum size of the window should be greater than the duty cycle of the appliances being dis-aggregated.

In the Chapter 7, Section 7.1.2.2, a brief study on the position of the sliding window around the time of analysis is proposed.

5.2 Meta-features

5.2.1 Principle

The problem of load identification being addressed here from the context of temporal classification, the issue is to convert raw data into a model that can be understood by established machine learning techniques. There are three broad approaches in the temporal classification domain :

1. Algorithms which deal specifically with temporal classification, for example, factorial hidden Markov models and sparse coding [Parson 2011, Labeeuw 2013].
2. Relational learning based techniques, like the recurrent neural networks.
3. Problem representation in a way that can be understood by propositional concept learners [Kadous 2002].

The work presented here is based on the third approach. Knowledge extraction for a specific representation of a problem is a technique of attribute construction applied to represent the underlying substructure of the training instances. In the temporal classification domain, these substructures are in the form of sub-events, defining for example a periodicity in the data [Dong 2012]. These sub-events become synthetic features, which are then fed to a propositional learner.

This concept also allows the inclusion of background and domain knowledge for temporal classification. The output of the learner can then be converted back to a human readable form for example as a decision tree.

5.2.2 Proposed Meta-features

One of the primary goals of this work is to define and use *meta-features* based on requirements of residential load identification. They are used to identify the substructures present in the aggregated power readings from the meter panel in order to identify the appliances signatures. The chosen meta-features are dedicated to the specific domain of high energy consuming loads in household buildings. Each of these features takes into account the different characteristics of appliances such as time of use, duration of use, trends of load, sequence of load, spike in load and correlation among appliances.

The main meta-features on which the identification algorithms are based are presented in the following subsections. Almost all of the meta-features are defined for the subsequence centered on the considered time of event t, except for the "hour of the day". The size of the sliding window at one moment of computation is $2N$ with $N = 5$ (refer to 5.1). Some of the presented meta-features are illustrated on a generic time window in the Figure 5.3.

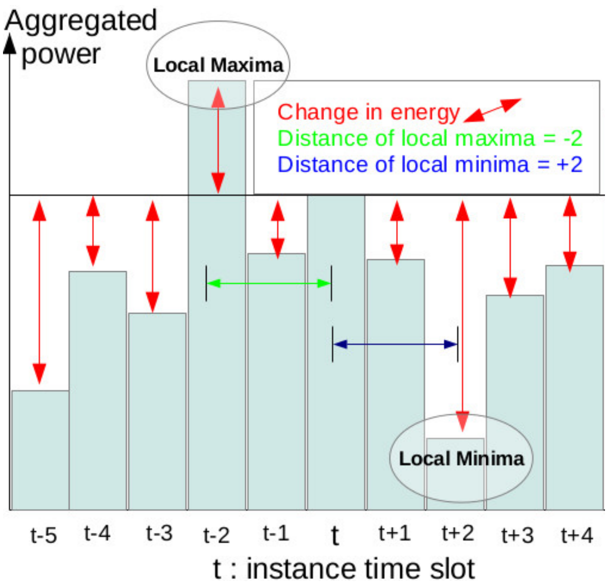
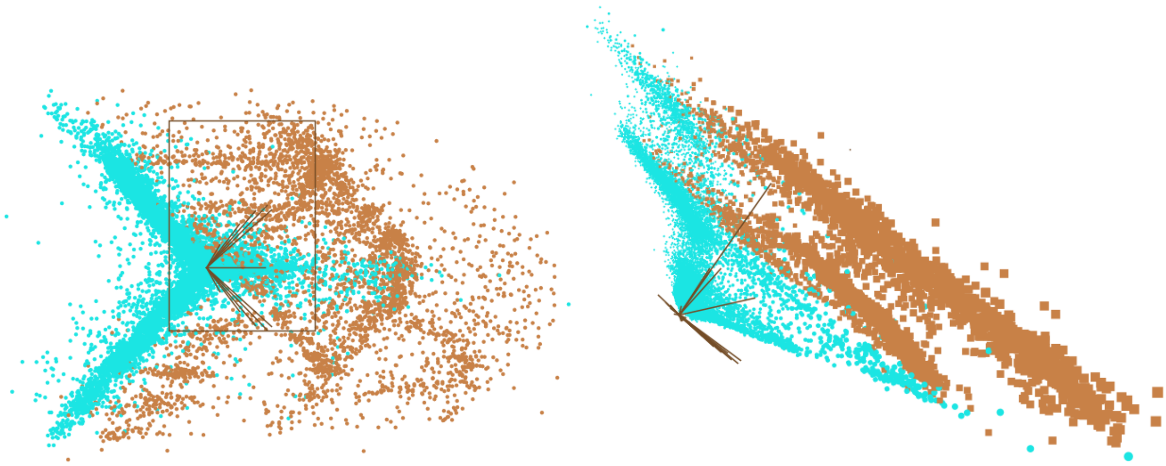


FIGURE 5.3 – Graphical definition of some of the main meta-features concerning the aggregated power measurements in a sliding window.

To visualize the impact of the use of meta-features on classification of states of applainces, a two dimensional projection plot (using PCA) of the water-heater ON/OFF states is proposed in the Figure 5.4. The figure 5.4a is generated using the energy consumption sequence as the feature, the figure 5.4b is generated using the meta-features as the feature. The ON state is represented by color brown and the color blue for the OFF state.

It can be clearly observed that using the meta-features as input of the classifier (in Figure 5.4b), the two states of the water heater can be considerably better separated than using only energy consumption information. Indeed, the process of the classifier algorithm can be easily understood graphically in two dimensions.

The task of the classification algorithm is to determine a line (for linear classifiers) or a non-linear curve (for non-linear classifiers) to best classify (i.e. distinguish between) the two classes. In can be intuitively observed by comparing the Figures 5.4a and 5.4b that after the use



(a) Based on the energy consumption sequence. (b) Using meta-features for classification.

FIGURE 5.4 – Two dimensional projection of the ON/OFF states of a water heater using PCA. The pixel size is proportional to the energy consumption.

of meta-features as input, the two classes are more separable. The classification error is thereby reduced.

In the following subsections, we will describe in a more detailed manner the major meta-features defined and tested in our work.

5.2.2.1 Hour of the day

The hour of the day $H(t)$ is a measurement of the hour of occurrence of an event. It is represented as a numeric value from 0 to 23 in the propositional learner, as described in equation 5.1.

$$H(t) = h; h \in [0, 23] \tag{5.1}$$

The impact of the hour of the day is illustrated in the Figure 5.5. In this figure, it can be observed that the appliance is primarily used during two periods during the day. Obviously, this feature plays an important role in representing the hourly temporal pattern in appliance usage.

5.2.2.2 Distances from the current event to the local maximum and local minimum

These two distances monitor the position, counted from the current time of event t, of the local maximum $d_M(t)$ and minimum $d_m(t)$ of energy consumption in the sliding window. This meta-feature provides to the classifier information on whether the current event is a local minimum, local maximum or neither. The displacement is measured as an integer value where "0" signifies whether the current event is a local minimum or maximum and "$+d$" or "$-d$" represents the distance (in time steps) to the local minimum and maximum in the subsequence.

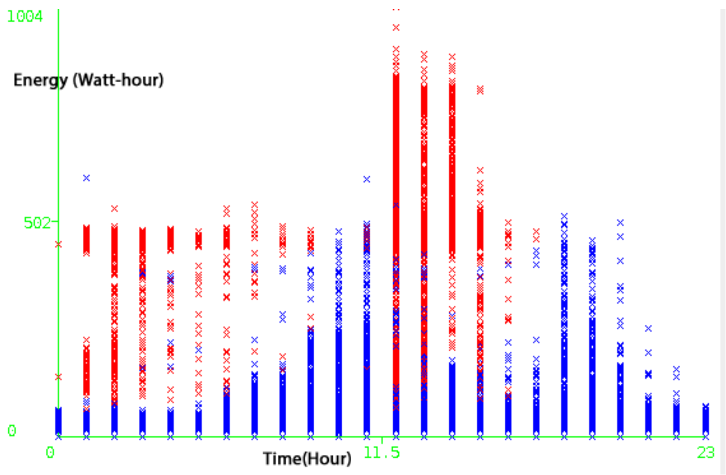


FIGURE 5.5 – Hourly usage of the water heater in one of the house of the IRISE database over one year. Two states are considered : OFF in blue and ON in red.

$$d_M \rightarrow \begin{cases} E(t + d_M) = E_{max} \\ E_{max} = \max\{E(t_i)\}; t_i \in [-N, N] \end{cases} \tag{5.2}$$

$$d_m \rightarrow \begin{cases} E(t + d_m) = E_{min} \\ E_{min} = \min\{E(t_i)\}; t_i \in [-N, N] \end{cases} \tag{5.3}$$

Where $E(t_i)$ is the energy consumed from t_{i-1} to t_i.

In the Figure 5.6 this meta-feature has been used to visualized the ON/OFF states of a water heater in one of the house of the IRISE database. The features are projected using the principle of PCA. As already shown in the Figure 5.4, the use of a meta-feature increases the two class separability compared to a trivial assessment.

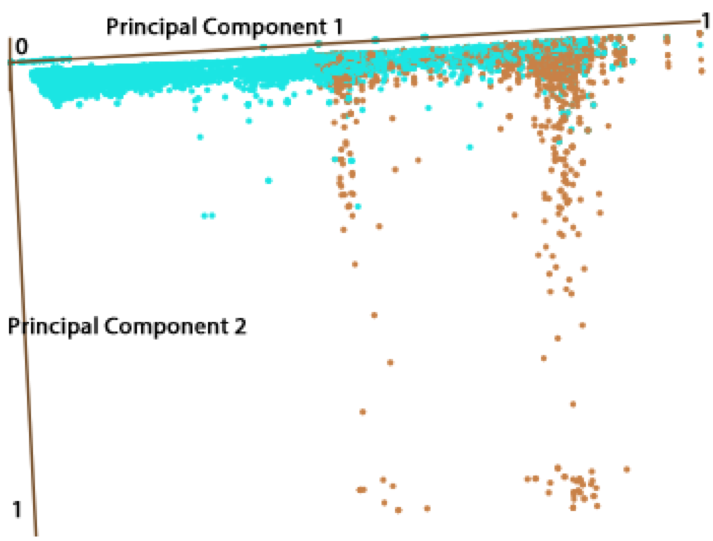


FIGURE 5.6 – Water heater ON/OFF states projected using PCA for the meta-feature called *distances from the current event to the local maximum and local minimum.*

5.2.2.3 Energy variation between time steps

This meta-feature takes into account the energy variations between the current time of event t and all other time steps in the considered sliding window, t_i.

$$E_v(t_i) = E(t) - E(t_i); t_i \in [-N+1, N] \tag{5.4}$$

This knowledge is directly representative of the appliance energy consumption and its variations in time, which is an indirect image of the load profile.

5.2.2.4 Gradient and Laplacian of energy consumption within the window

For this meta-feature, the gradient ∇, laplacian Δ and gradient ratio ∇_r are evaluated around the current time of event t. These rates of change within the sliding window allow the classification algorithm to identify trends in energy consumption. The gradient ratio is the ratio between the gradient and the aggregated power value at any given time instance t_i of the sliding window.

$$\left.\begin{aligned} \nabla(t_i) &= E(t_i) - E(t_i - 1) \\ \Delta(t_i) &= \nabla(t_i) - \nabla(t_i - 1) \\ \nabla_r(t_i) &= \nabla(t_i)/E(t_i) \end{aligned}\right\} t_i \in [-N+1, N] \tag{5.5}$$

These equations are used to identify different energy spikes or edges and also take into account the base energy level from which the spikes occur. At low sampling rates, two appliances may present similar edges but different base energy level makes them differentiable.

5.2.2.5 Mean and standard deviation

This meta-feature is used to add global statistics on the nature of all the other meta-features within the sliding window. The mean and standard deviation are computed in the considered sliding window for the aggregated power E, for the change of energy from previous to current state E_v, and for the first derivatives of energy ∇.

$$\left.\begin{aligned} \overline{E}(t) &= \frac{1}{2N}\left(\sum_{t_i=-N}^{N} E(t_i)\right) \\ \overline{E_v}(t) &= \frac{1}{2N}\left(\sum_{t_i=-N}^{N} E_v(t_i)\right) \\ \overline{\nabla}(t) &= \frac{1}{2N}\left(\sum_{t_i=-N}^{N} \nabla(t_i)\right) \end{aligned}\right\} t_i \in [-N+1, N] \tag{5.6}$$

Conclusion

The work of Non-intrusive load monitoring can be studied as a temporal classification problem. This problem encompasses many domains and thus requires a detailed understanding and a broader perspective. Therefore, the relevant domains for the problem at the low rate are discussed. The raw data is transformed using sliding window into a form that could be directly learned by a propositional learner. The input features are further enriched by generating temporal meta-features. The various stages of data transformation using meta-features is explained both logically and visually. The novel meta-features choices are explained and their relevance emphasized. It is also observed visually that the use of domain specific meta-features resulted in a better disaggregation capability for the subsequent classifier.

Chapitre 6

Classification algorithms

6.1 Multi-label classification 58
6.1.1 Binary Relevance 59
6.1.2 Label Powerset 59
6.1.2.1 Rakel algorithm 60
6.1.3 The Decision Tree Learner 60
6.1.3.1 Disparity measurements 60
6.1.3.2 Decision process 61
6.1.4 The Support Vector Machine algorithm 62
6.2 The K-Nearest Neighbours classifier 62
6.2.1 Euclidean Distance 63
6.2.2 Dynamic Time Warping 63
6.2.3 Temporal Correlation 64
6.2.3.1 Value and behavior based metric 65
6.2.4 Multi-label k-nearest neighbour 65
6.3 Hidden Markov Model based classification 66
6.4 Imbalance in the data-set 67
6.4.1 Instance sampling during training 68
6.5 Evaluation Strategies 68
6.5.1 Indicators of classifier performance 68
6.5.2 Evaluation procedures 70

Introduction

The problem is represented in the form of a standard propositional learner, after the sliding window process and meta-feature generation discussed in the previous chapter. The subsequent remaining task is to train or learn a function which maps the input-output relationship. As a reminder, in our work the input is constituted by the energy readings and the meta-features and the output is the appliances states.

It is further observed that the outputs (appliances) are correlated to their usage. In this chapter, different classification techniques are explored to converge to a choice of suitable classification techniques which can model the requirement of the classification system. Finally, the difficulties of the imbalanced dataset and relevant evaluation strategies are discussed.

6.1 Multi-label classification

The identification of the state of loads is based on multi-label classification techniques. These techniques are frequently used in the field of information theory and present many advantages over single-label classification [Tsoumakas 2011, Tsoumakas 2007]. A classification learner approximates a function, mapping a vector into labels by looking at input-output examples of this function. The features x_i and the target class Y come in record of the form :

$$(x, Y) = (x_1, x_2, x_3, \dots, x_n, Y) \tag{6.1}$$

For multi-label classification : $Y = \{0, 1\}^L$ where L is the number of appliances. In multi-label classification problems, multiple target labels are assigned to each instance. In this work, a function is built which maps inputs x_i to an output vector Y rather than a scalar output which is the case for single-label classification. Given a data-set of labelled instances, classification algorithms seek relations that will correctly predict the class of future unlabelled instances Y', from future features x', where :

$$Y' = f(x') = \begin{cases} 0 & \text{if the appliance is OFF} \\ 1 & \text{if the appliance is ON} \end{cases} \tag{6.2}$$

The fact that multi-label classification takes into account the interdependence among labels (in our case the appliances) is the main reason why this work is based on these classifiers. For example, generally the clothes drier is used after the washing machine. Multi-label classifier models are built on *meta-features* as inputs which are computed on the sub-sequences C_k. The appliance states are the output classes in this case. The time stamped data also allow taking into account temporal information for load identification (temporal patterns).

There are two broad approaches in handling with multi-label classification algorithm. One is by way of problem transformation where a multi-label problem is transformed into one or more single-label problems and then a state of the art classification algorithm such as decision tree or support vector machines is used. Another is to modify an existing single-label algorithm directly for the purpose of multi-label classification (algorithm adaptation method e.g ML-kNN). Two multi-label problem transformation and three different classification algorithms are implemented and compared in this work. A Hidden Markov Model is also implemented for comparison. A brief description of these algorithms follows.

Table 6.1 – Example of multi-label classification

Attributes				classes of labels			
Attr.1	Attr.2	...	Attr.N	Lbl.1	Lbl.2	...	Lbl.L
Val.1	Val.1	...	Val.1	Cl.1	Cl.2	...	Cl.1
Val.2	Val.2	...	Val.2	Cl.1	Cl.2	...	Cl.2
Val.3	Val.3	...	Val.3	Cl.2	Cl.1	...	Cl.1

Table 6.1 illustrates an example of multi-label classification problem. In this table, for three

arbitrary instances, all combination of L labels classes and N attributes values are monitored. In this table, labels can only have two classes or states.

In the following subsections, we will describe different classifiers, and compare their qualities.

6.1.1 Binary Relevance, Br

The Br transformation is a method of problem transformation that learns separately single-label binary models for each classes or label [Tsoumakas 2011, Tsoumakas 2007]. It transforms the original data into single label data-sets that contain all the examples of the original data-set. From the multi-label classification example presented in Table 6.1, the Br algorithm will build the same number of tables as there are labels, each one of them having all the attributes and only one label. Table 6.2 is generated from Table 6.1 using a Br transformation.

Table 6.2 – One of the tables obtained with a Br transformation from the Table 6.1

Attributes				Classes of label
Attr.1	Attr.2	...	Attr.N	Label i
Val.1	Val.1	...	Val.1	Cl.1
Val.2	Val.2	...	Val.2	Cl.1
Val.3	Val.3	...	Val.3	Cl.2

6.1.2 Label Powerset, Lp

The Lp transformation considers each different set of labels that exist in the multi-label data-set as one single label [Tsoumakas 2011, Tsoumakas 2007]. Unlike the Br classifier, the Lp algorithm learns using only one single classifier consisting of the number of classes times the number of labels in the original multi-label problem. The primary advantage of using this transformation is that it takes into account appliances correlations. The primary drawback is the computation cost compared to the Br transformation.

As an illustration, an example of table that the Lp transformation would give from the multi-label classification presented in the Table 6.1 is proposed in the Table 6.3.

Table 6.3 – Example of Lp transformation from the Table 6.1

Attributes				Classes of label
Attr.1	Attr.2	...	Attr.N	New label
Val.1	Val.1	...	Val.1	Lbl.1(Cl.1) $\wedge$ Lbl.2(Cl.2) $\wedge$ ...$\wedge$ lbl.L(Cl.1)
Val.2	Val.2	...	Val.2	Lbl.1(Cl.1) $\wedge$ Lbl.2(Cl.2) $\wedge$ ...$\wedge$ lbl.L(Cl.2)
Val.3	Val.3	...	Val.3	Lbl.1(Cl.2) $\wedge$ Lbl.2(Cl.1) $\wedge$ ...$\wedge$ lbl.L(Cl.1)

6.1.2.1 Rakel algorithm

Label Powerset algorithm discussed previously suffers from run time complexity. As the number of appliances increases the complexity increases exponentially. The Rakel algorithm uses an ensemble of subset of Label Powerset (LP) transformation ; it randomly selects a subset of labels and tries to lean the model. The ensemble combination is accompanied by "thresholding" the average zero-one decision of each model per considered label [Tsoumakas 2010]. Its performance is generally similar to Label Powerset but with reduced complexity.

6.1.3 The Decision Tree Learner, DTL

The decision tree consists of nodes where a logical decision has to be made. Branches are connected according to the result of these decisions. For each node of the tree, one attribute of the data is selected that most effectively splits its set of samples into subsets enriched in one class or the other. Following a path of nodes and branches constitute a sequence through a decision tree that reaches to a final decision regarding a specific appliance (in our case, the ON or OFF state).

The DTL algorithms represent one of the preferred choices for load classification as described in [Quinlan 1986]. Indeed, decision trees are rule based and the built model is easy to visualize. The Figure 6.1 proposes an example of DTL visualization learned on a typical appliance in one of the houses of the IRISE database.

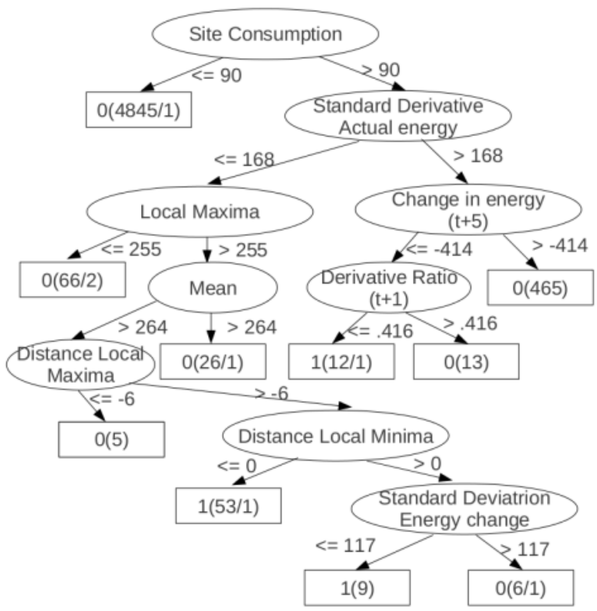


FIGURE 6.1 – Construction of a decision tree, applied on the application "electric oven".

6.1.3.1 Disparity measurements

The impact of the meta-features into a learning system can be well understood by the use of the *disparity measure*. In this work, the disparity measure is the gain ratio of the decision tree learner, as described in the previous section (section 6.1.3).

The position of the feature in the tree gives an idea about the significance of an attribute. The closer the feature is to the root, the higher its gain ratio is and thereby its significance. An example of typical decision trees obtained during identification is proposed in the Figure 6.1.

A good quantitative measure of the significance of an attribute is a statistical property called *information gain*. It measures how well a given attribute separates the training examples according to their target classification. This measure is used to select among the candidates attributes at each step while the decision tree is growing. The attribute presenting the highest normalized information gain is chosen to make the decision.

The metric used in practice is the gain ratio which corrects the information gain by taking the intrinsic information of a split into account. Then, the algorithm applies it recursively on the sub-lists. Finally, the impact of the meta-features into a learning system can be well understood by the use of the *disparity measure*. In this work, the J48 implementation of *C4.5* algorithm is used [Weka 2014] and the parameter is optimized using parameter section algorithm during training.

The DTL usually leads to a good understanding of the significant features for each appliance. Based on the disparity measurement, the attribute with the highest normalized information gain is chosen as the root of the decision tree. Information gain is measured in bits and is given a probability distribution, the information required to predict an event is the distribution's entropy, given by :

$$S(p_1, p_2, \dots, p_n) = -p_1 \log(p_1) - p_2 \log(p_2) - \cdots - p_n \log(p_n) \tag{6.3}$$

6.1.3.2 Decision process

The *leaf nodes* in the decision tree gives the number of instances correctly classified by the built model. For example, in the Figure 6.1, "$0(x, y)$" means that x is the number of instances correctly classified and that y is the number of instances incorrectly classified. Also, the value "0" means that the considered appliance is OFF and "1" means ON.

This can be corroborated by looking at the increasing number of correctly predicted cases when approaching the root of the decision tree. Indeed, the higher in the decision tree an attribute is, the more significant it is. The number of correctly classified cases should then be following accordingly.

As it can be seen in the Figure 6.1, this decision process is easy to apprehend as its representation is based on logical decisions. It is a way to validate our proposed meta-features. In fact, the main meta-features proposed in the section 5.2 of the Chapter 5 have appeared in the decision trees built during the identification tests.

6.1.4 The Support Vector Machine algorithm, (SVM)

The SVM algorithm is a powerful tool for data classification described in [Onoda 2000]. The first major step of a SVM classification is to build a decision plane that separates a set of objects with different class memberships. It guarantees the best function to distinguish between members of classes by maximizing the margin between them. The maximal margin hyper-planes allow the best generalization abilities and thus the best classification performances on the training dataset. This procedure requires finding the solution of the following optimization problem :

$$\min_{\mathrm{w},b,\xi} \left(\frac{1}{2}\mathrm{w}^T\mathrm{w} + C\sum_{i=1}^{l} \xi_i \right) \tag{6.4}$$

$$\text{subject to } \begin{cases} y_i\left(\mathrm{w}^T\phi(x_i)+b\right) \geq 1-\xi_i \\ \xi_i \geq 0 \end{cases} \tag{6.5}$$

with l the total number of sub-sequences, w the normal vector of the hyper-plane, b the offset of the hyper-plane, C the penalty parameter of the error term ξ and ϕ the kernel function.

The second major step is to choose the kernel function of the algorithm. The Radial Basis function is preferred over others in this work. For two groups i and j, the training vectors x_i and x_j are mapped to a higher dimensional space by the kernel function ϕ defined as :

$$\begin{cases} \mathrm{K}(x_i, x_j) & \equiv \phi(x_i)^T\phi(x_j) \\ \mathrm{K}(x_i, x_j) & = \exp\left(-\gamma\|x_i - x_j\|^2\right); \gamma > 0 \end{cases} \tag{6.6}$$

where γ is a parameter of the kernel. In our work, a grid-search has been conducted on the parameters C et γ using cross-validation. Without being used here, other kernels exist for the SVM algorithm :

Linear kernel : $\mathrm{K}(x_i, x_j) = x_i^T x_j$
Polynomial kernel : $\mathrm{K}(x_i, x_j) = (\gamma.x_i^T x_j + r)^d, \gamma > 0$
Sigmoid kernel : $\mathrm{K}(x_i, x_j) = \tanh(\gamma.x_i^T x_j + r)$

The SVM algorithm is computationally more expensive than rule based algorithms such as DTL. In this work, the Sequential Minimal Optimization (SMO) implementation of [Weka 2014] is used with a grid search for parameter optimization during training.

6.2 The K-Nearest Neighbours classifier, KNN

The KNN classifier is an instance based learning method where the classification function is approximated by a majority vote of the neighbors using a distance metric. K is the number of neighbors which is calculated using cross-validation. The function is approximated locally and all computation is deferred until classification. Typically, the Euclidean distance is used as distance function. But in this work, time series metrics has been preferred.

The mechanism of KNN is quite simple. For any new data instance, the attributes of the new case is compared with all the previously seen cases or instances in the training database. The comparison is typically based on a distance measurements. The nearest instances or cases in the training database are evaluated based on the distance metric. The new instance is assigned to the class of the majority of neighboring instances (process of classification). Mathematically, for any instance x_i in the database of size $n \times x$, the distance is expressed as :

$$d(x_i) = \min_j d(x_i, x) \text{ with } j \in \{1, \cdots, n\} \tag{6.7}$$

The key point here is the use of the proper distance metric. The default metric is the Euclidean distance, which is calculated on the normalized value of the attributes. In our case, other possibilities are explored : the temporal distance metrics. These metrics and their interest are presented in the following subsection.

6.2.1 Euclidean Distance

The Euclidean distance is a standard metric to calculate the distance between two points. It is given by the following equation :

$$d_E(X, Y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \tag{6.8}$$

where n is the total number of points. This is a static measure and is not able to account for the temporal relation between the data points.

6.2.2 Dynamic Time Warping

Dynamic time warping (DTW) is a template-matching recognition method based around a dynamic programming algorithm [Douzal-Chouakria 2012]. They are primarily used to determine similarity between two temporal sequences which may vary in time or speed.

The basic aim of the dynamic time warping algorithm is to align the time axes of a sampled time series and a template, in order to minimize some distance measures. The time axis of either series is stretched or compressed to achieve the best possible fit, allowing a single template to match a range of similar patterns in the data.

Both the template and the time series consist of sequences of data points over a time interval, which does not need to be the same for both sequences. The dynamic time warping problem is stated as follows :

Let $X = (u_i, \cdots, p)$ and $Y = (v_1, \cdots, q)$ be two time series of length p and q.

Construct a warp path :

$$D = d_1, d_2, \cdots, d_k \quad max(p, q) \leq K < p + q \tag{6.9}$$

k is the length of the warp path and the k'th element of the warp path is $d(i, j)$. Here, i is an index from time series X, and j is an index from time series Y. The warp path must start at the beginning of each time series at $D_1 = (1, 1)$ and finish at the end of both time series at $D_k = (p, q)$.. So both the ends of the time-series is bounded. The warp is also monotonically increasing which could be formally represented as :

$$D_k = (i, j) \quad D_{k+1} = (i', j') \quad ; \quad i \leq i' \leq i + 1 \quad ; \quad j \leq j' \leq j + 1 \tag{6.10}$$

The optimal warp path is the warp path is the minimum-distance warp path, where the distance of a warp path W is :

$$Dist(W) = \sum_{k=1}^{k} = KDist(w_{ki}, w_{kj}) \tag{6.11}$$

Dist(D) is the distance (typically Euclidean distance) of warp path D, and $Dist(w_{ki}, w_{kj})$ is the distance between the two data point indexes (one from X and one from Y) in the k'th element of the warp path.

The above cost function involves the differences between the aligned values of two time series, without taking into account the values neighborhoods. A visual illustration of the distance between two time series having similar values-based characteristics is proposed in Figure 6.2.

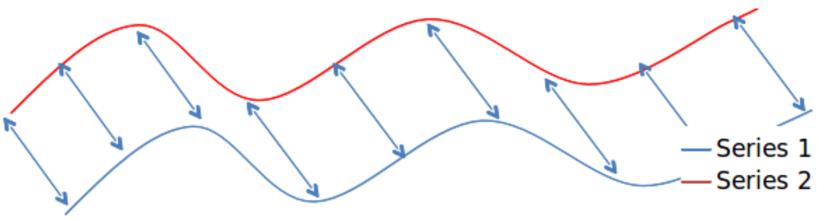


FIGURE 6.2 – Two time Series having similar values-based characteristics

6.2.3 Temporal Correlation

Two time series X and Y are considered similar in behavior if, during any observed period $[t_i, t_{i+1}]$, they increase or decrease simultaneously with the same growth rate. In contrast, they are considered to be opposite in behavior if, during any observed period $[t_i, t_{i+1}]$ in which X increases, Y decreases and (vice-versa) with the same growth rate (in absolute value).

Let $X = (u_{i,\cdots,p})$ and $Y = (v_{1,\cdots,q})$ be two time series of length p and q. The Pearson correlation coefficient has been used as a behavior proximity measure between signals. An equivalent formula for the correlation coefficient relying on pairwise values differences is shown below :

$$Cor(X, Y) = \frac{\sum_{i,i'}(u_i - u_{i'})(v_i - v_{i'})}{\sqrt{\sum_{i,i'}(u_i - u_{i'})^2}\sqrt{\sum_{i,i'}(v_i - v_{i'})^2}} \tag{6.12}$$

It can be seen that the correlation coefficient assumes the independence of data as based on the difference between all of the pairs of values observed at $[t_i, t_{i'}]$. In contrast, the behavior proximity needs only to capture how time series behave at $[t_i, t_{i+1}]$. Thus, the correlation coefficient is biased by all of the remaining pairs of values observed at $[t_i, t_{i'}]$ with $|i - i'| > 1$.

For temporal data, a variant of the Pearson correlation involving first-order differences is :

$$Cort(X,Y) = \frac{\sum_i (u_{i+1} - u_i)(v_{i+1} - v_i)}{\sqrt{\sum_i (u_{i+1} - u_i)^2}\sqrt{\sum_i (v_{i+1} - v_i)^2}} \tag{6.13}$$

with $Cort(X,Y)$ belonging to $[-1;1]$. $Cort(X,Y) = 1$ indicates a similar behavior and $Cort(X,Y) = -1$ an opposite behavior between X and Y.

A visual illustration of the distance between two time series having similar behavior-based characteristics is proposed in Figure 6.3.

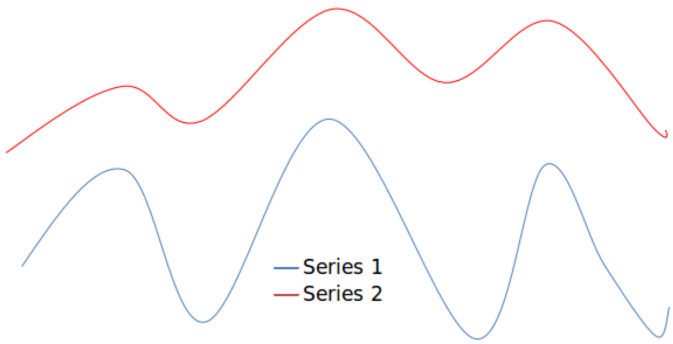


FIGURE 6.3 – Two time Series having similar behavior-based characteristics (Temporal Correlation, TC)

6.2.3.1 Value and behavior based metric

A last possibility has been proposed in [Douzal-Chouakria 2009] : a proximity measure covering both the value and behavior based cost function. A parameter k is used for modulation between the two original metrics.

$$c_k(r) = \frac{2}{1 + \exp(k.Cort(S_1, S_2))}.d_{Dtw}, 0 \leq k \tag{6.14}$$

In our work, values from 0 to 5 have been used for the parameter k for fine tuning the modulation effect.

6.2.4 Multi-label k-nearest neighbour (ML-KNN)

ML-KNN is the multi-label implementation of the k-nearest neighbor algorithm for single-label classification [Zhang 2007]. It is a direct implementation for multi-label problem so a single-label base classifier is not required. It works on the principle that for every test instance, its k nearest neighbors in the training set are identified. Then, according to statistical information gained from the label sets of these neighboring instances (i.e. the number of neighboring instances belonging to each possible class), a maximum a-posteriori principle is used to determine the label set for the test instance. ML-KNN is a learning approach which is different from the rule and function based approaches discussed above.

6.3 Hidden Markov Model based classification, HMM

HMM is a standard parametric statistical modelling technique used extensively in the field of signal processing to maximize the probability that a series could have been generated by a model. A Markov model is a sequence of random variables in which each variable is dependent upon only the variable immediately preceding it. A hidden Markov model (HMM) is a Markov model in which the sequence is made up of discrete state variables. An associated continuous variable x is observed for each discrete variable z, in our case the state of the considered appliance.

There are three major applications of HMM in real life problems : the evaluation problems, the decoding problems and the learning problems. In this application the learning problem approach is used. In the learning approach, given a model $\lambda(A, B, \pi)$ and a set of observations x, the problem is to find the optimal model parameters that maximize the probability that the model produced the proper observations. The solution to this problem provides the means to train a model to recognize a particular sequence [Rabiner 1986].

Iterative procedures can be used to solve this problem. One of the popular algorithm for that matter is the forward-backward algorithm, also known as the Baum-Welch or expectation-maximization method [Rabiner 1986].

In our work, the input is the time series sub-sequences of energy readings from the power meter. The number of states is optimally determined through cross-validation. It can be observed in the Figure 6.4 that the observed continuous variable in our case is the total energy consumption at the smart meter and the hidden discrete variables are the appliance states.

The three model parameters which are evaluated are :

$A_{i,j}$ **:** The state transitional probability, of being at state A_i and making a transition to A_j in the next time instance. Mathematically, $A_{i,j} = P(\{z_{t+1} = j\}|\{z_t = i\})$. In the implementation, the state transitional matrix is initialized using a K-means algorithm.

$B_{x|z}$ **:** The emission probability for having the observation x_t being at state z_t. For continuous variables, a Gaussian distribution is generally used.

π_z **:** The probability of each state of the discrete hidden Markov model.

The HMM is implemented in the following set of sequences :

1. Pre-processing and generating the time-series sub-sequences.
2. Initializing and re-estimating the HMM parameters from the training data.
3. Evaluating the output class based on learned model parameters on unseen testing examples.

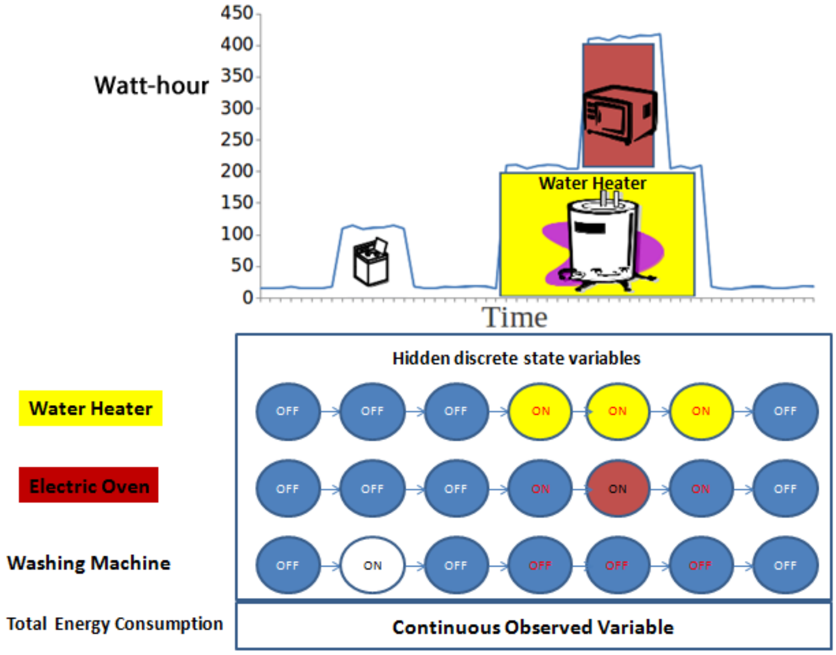


FIGURE 6.4 – The continuous and discrete variables for the Hidden Markov Model applied to load identification.

6.4 Imbalance in the data-set

A data-set is imbalanced if the classification categories are not approximately equally represented. In our application, the number of instances where an appliance is in the ON state is considerably low compared to the number of OFF state instances. Such imbalance in the data-set creates many challenges to the learning algorithm, as illustrated in the Figure 6.5. The two major challenges faced in our application are during the training and the evaluation of the used measures.

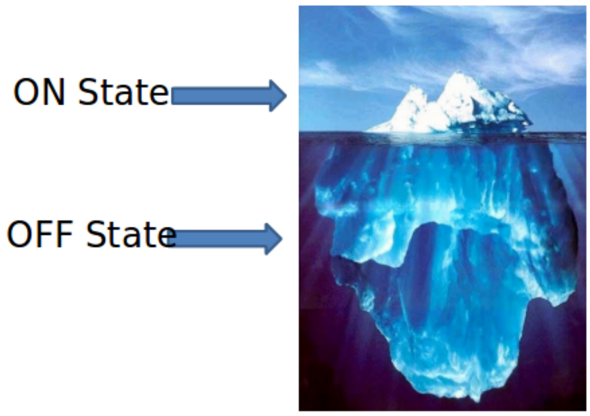


FIGURE 6.5 – Iceberg to showcase the imbalance in the dataset

6.4.1 Instance sampling during training

To deal with imbalance in the data-set many sub-sampling and over-sampling techniques can be found in the literature [Chawla 2011]. In sub-sampling, the number of instances of the majority class is reduced during training using some selection criteria whereas in over-sampling the number of instances of the minority class is increased artificially during training to reduce the imbalance in the data-set. In this work one sub-sampling (spread sub-sample) and over-sampling (SMOTE) techniques were implemented. Experimentally sub-sampling performed better for this domain and was used during training for enhanced performance.

The spread sub-sampling randomly selects data points (instances) from the original data reducing the number of instances of the majority class and making the dataset more balanced. Using this method in our case results in a significant reduction in training instances. Heuristically, the 2 :1 ratio of the classes were chosen, where the majority class is twice the number of minority class.

The Synthetic Minority Oversampling Technique (SMOTE) is a oversampling technique where the number of minority classes are increased. The algorithm uses the nearest neighbor algorithm to artificially generate features. This is done by introducing synthetic examples along the line segments joining any or all of the k minority class neighbors. Depending on the percentage of oversampling the nearest neighbors are randomly chosen. In the implementation a 200 percent oversampling was done. Initial results indicate that the sub-sampling method works better in our case the over-sampling method.

6.5 Evaluation Strategies

6.5.1 Indicators of classifier performance

Given a dataset of labeled instances, supervised machine learning algorithms seek an hypothesis that will correctly predict the class of future unlabeled instances. As already discussed in the Chapter 6 in Section 6.5, in order to compare structures of predictors, we need indicators that will give a quantitative way of assessing the classifier performances. While comparing these indicators values, the best predictor can be found for a given appliance.

In order to properly define the performance indexes of the classification algorithms used for prediction, we introduce the *confusion matrix* [Kohavi 1998].

A confusion matrix contains information about the actual and the predicted results obtained by a classification system. The performances of such systems are commonly evaluated using the data contained in this confusion matrix. The table 6.4 shows the confusion matrix for a two-class classifier. The classes that can be predicted are “positive" or “negative" instances, which in this case signifies the appliance consumes or does not consume energy.

In the context of this study, the entries defined in the confusion matrix reported in the Table 6.4 have the following meaning :

a : is the number of correct predictions where an instance is negative,

Table 6.4 – Confusion matrix

		Predicted	
		Negative	Positive
Actual	Negative	a	b
	Positive	c	d

b : is the number of incorrect predictions where an instance is positive,

c : is the number of incorrect of predictions where an instance negative,

d : is the number of correct predictions where an instance is positive.

Several standard terms have been defined for this 2 class matrix :

The accuracy : is the proportion of the total number of predictions that were correct. It is determined using the ratio $AC = (a+d)/(a+b+c+d)$. It is the percentage of cases where the predicted energy state (On or Off) is correct for an appliance.

The true positive rate (recall) : is the proportion of positive cases that were correctly identified, as calculated using the equation $TP = d/(c+d)$. Represents the ratio between the predicted positives states of the appliances (On) and the total number of correct positives states of the appliances.

the precision : is the proportion of the predicted positive cases that were correct, as calculated using the equation $P = d/(b+d)$. Represents the fraction of the positives states (On) of the appliances correctly predicted.

There are multiple labels or appliances present when we represent the problem as a multi label classification problem. To evaluate such a model we need go beyond accuracy and look into other evaluation measures which deal with the appliances all together. For this reason, the confidence of the predictions is monitored in our work using tools commonly used in information theory [Tsoumakas 2011, Tsoumakas 2007] with use the measures defined below.

F-measure : Is taken to be the weighted harmonic mean of precision and recall. In this work, a special case of the general F-measure definition has been used, where both are contributions of the precision and the recall are equally weighted in the F-measure coefficient.

Receiver operating characteristic - Area Under Curve (Auc) : Is a graphical plot of the true to false positive rate at various threshold settings of the classification algorithm. The Auc score is given on a scale from 0 to 1, where a score of 1.0 indicates perfect classification and a score below 0.5 shows a quasi-random guessing.

Subset accuracy : is defined as the percentage of cases where the disaggregation of the load into its constituent appliances (under investigation) are correct. This is also a measure which gives an idea about how well the energy disambiguation is working.

These measures in the multi-label classification require some additional metrics than those used in traditional single-label classification. The Micro and Macro averaged scores are discussed for the same. Micro-averaged values are calculated by constructing a global contingency table and then calculating precision and recall using these sums. Macro averaged scores are calculated by first calculating the measure(s) for each appliance and then taking the average of them. In this case the Macro averaged results are of more importance, as the data is sparse the micro average scores are generally high.

All these qualification coefficient will be used in the results chapter (Chapter 7, 2 and 10) for proper interpretation of the load identifications and prediction.

6.5.2 Evaluation procedures

There are two standardized methods to evaluate the performance of a classification algorithm. They are cross-validation and hold-out data analysis. In a 10 fold cross validation, the data of size n is divided in $n/10$ sets. The classifier is trained on nine sets of the dataset and tested on the remaining one. The process is repeated 10 times and the average measure is taken. This is standard method for comparing algorithms or evaluating parameters based on a measure.

The hold-out data analysis is using a portion of the data for training and the remaining for testing. In our application both this techniques are used depending on the requirement.

Summary

The method of multi-label classification is explained and the relevance to the current problem is highlighted. The major advantage of multi-label classification is the consideration of interclass correlation. The various problem transformation techniques (BR, LP) are also explained. The base leaner such as Decision Tree Learner (DTL) and Support Vector Machines (SVM) are summarized and the chosen configuration is discussed. The base learners represent different class of learners, the decision tree (C4.5) represent a rule based learner, the SVM represents a function based learner and the KNN is an instance based learner. The KNN algorithm using time series metrics is also discussed. The use of temporal distance measure is highlighted using both value and behavior based metrics. The Hidden Markov Model is used as a sequence learner in many application, a discrete HMM model suitable for the problem is shown. Finally, the imbalance of the data for such application is observed and remedies discussed, the evaluation of these imbalance and multi-label dataset is also summarized.

Chapitre 7

Load Monitoring Results

7.1 Time-related observations . . . 72

7.1.1 Usage of the loads in time . . . 72

7.1.1.1 Time-line . . . 72

7.1.1.2 Frequency of usage . . . 72

7.1.2 Classification algorithm and time analysis . . . 73

7.1.2.1 Time of training . . . 73

7.1.2.2 Sliding window . . . 73

7.2 Energy variation between time steps . . . 74

7.3 Identifying the loads . . . 75

7.4 Load identification for defined categories of consumers . . . 76

7.4.1 Considering the first category of houses . . . 76

7.4.2 Considering the second category of houses . . . 78

7.4.3 Considering the third category of houses . . . 79

7.5 Sensitivity to the sampling rate . . . 81

7.6 Comparison with standard non-intrusive load identification methods . 82

7.6.1 Hidden Markov Model . . . 82

7.6.2 Knn Distance based measures . . . 82

7.7 Discussion . . . 84

Introduction

In this chapter, various classification techniques discussed in the previous chapters are applied and the results are presented. Various aspect of the technique is explored, like the parameters selection, such as the orientation of the sliding window, the categories of houses and the sensitivity to sampling rate. Comparison to state of the art classification techniques are conducted. The used classifiers represent different categories of classifiers. Each category of classifier has its own advantages and disadvantages, the work further tries to observe the implication of the use of different categories classifiers and highlight the advantages and disadvantages. Finally, the use of temporal distance based techniques is illustrated, the results are compared for value and behavior based metrics. The observation from the results is discussed and further conclusion are drawn.

7.1 Time-related observations

In this section, the various temporal aspect of the model is explored. To have an overview on identification performance, in the section 7.1.1, the comparison of identification performance with actual case is observed. It is observed both in real time scale for the year and also in terms of number of usage at a particular hour.

7.1.1 Usage of the loads in time

7.1.1.1 Time-line

A way of analyzing the energy consumption is by visualizing its repartition of over a full year. As an illustration, the Figure 7.1 shows the time-line for an electric cooking device which is tested for a year using the Rakel algorithm on one house.

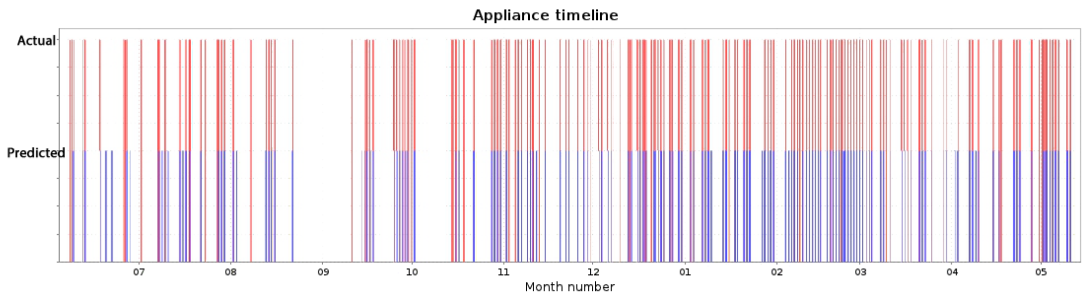


FIGURE 7.1 – Hourly time-line diagram, electric cooking (oven + hotplates). The red lines are the identified states of the load and the blue lines are the states from the readings database.

The time-line gives an overall idea about the usage of the appliance over the year and the months in which the usage is more intensive than in the others. In this figure, the red lines represent the ON states after the identification and the blue lines represent the ON state from the database (true state). When there is no possible superposition between both time series, it means that there is an error in the identification of the state of the load. It can be observed that the appliance has been identified with a reasonable amount of accuracy in this case. The results also highlight the period during which the major identification error occurs which needs to be further analysed.

7.1.1.2 Frequency of usage

The corresponding frequency of usage in terms of hours of the day is proposed in the Figure 7.2. In this figure the hourly frequency of usage represents the number of times the appliance was in its ON state in the considered 60 minute, divided in 6 time-slots of 10 minutes each. So if the appliance was ON for the whole hour the corresponding count will be 6 for that hour. This "frequency" is calculated for the length of the tested database. This predicted frequency chart may be provided to the inhabitants along with a more efficient energy management plan.

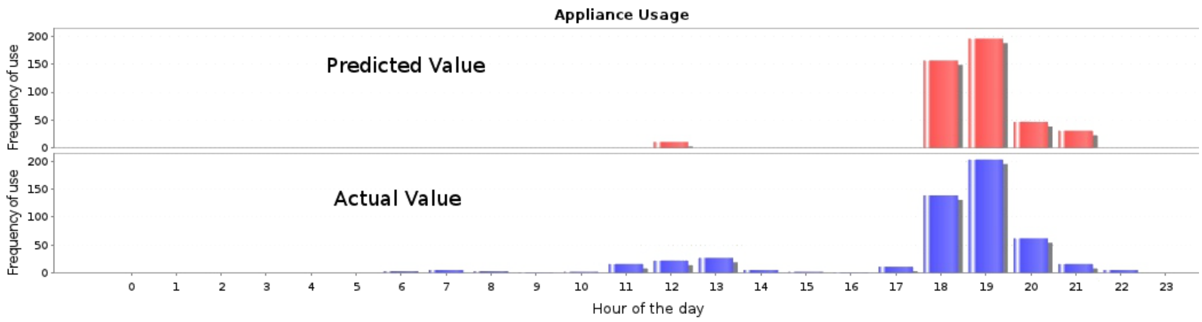


Figure 7.2 – Hourly frequency of usage diagram for an electric cooker (oven + hotplates).

7.1.2 Classification algorithm and time analysis

7.1.2.1 Time of training

The classification algorithm is trained on a certain portion of the database, and uses the rest for validation (testing). This time of testing is representative of the time inhabitants can obtain the result without fresh training. Therefore it is interesting to play with that duration in order to assess the sensitivity of the results to that testing period. In the Figure 7.3 a time based account of the classification results is shown for one random house from the Irise.

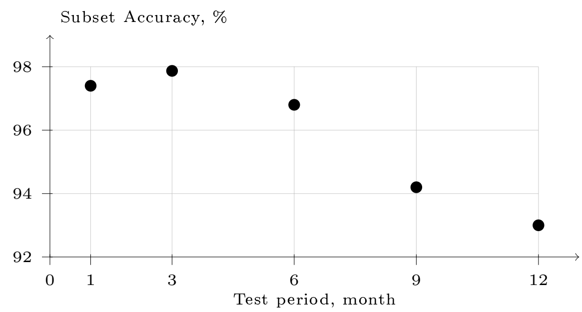


Figure 7.3 – Accuracy as testing time is increased.

If classification is conducted on a smaller period of time, the accuracy is much higher compared to longer period of testing. This fact may be attributed to multiple factors among which are the seasonal variations of usage of an appliance and it being operated in a different state than the one seen in the past. So, a testing period within 6-months is ideal after which for better results fresh training needs to be done.

7.1.2.2 Sliding window

When considering algorithms based on sliding windows process, three possible orientations can be chosen as they are presented in the Figure 7.4. The position of the sliding window can be balanced around the time of analysis, or left or right oriented.

The multi-label classification algorithm Rakel is used (as presented in the previous chapter)

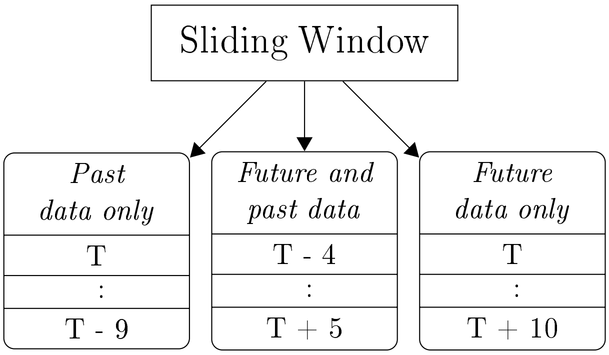


FIGURE 7.4 – Three possible choices for the sliding window position

[Tsoumakas 2010] with the decision tree (DTL) as the base learner. The table 7.1 shows the difference of efficiency of this classification algorithm on three houses from the IRISE database. For an explanation on the qualifying coefficient, please refer to the Section 6.5 of the Chapter 6.

TABLE 7.1 – Selection of sliding window orientations.

Test house	Measure	left	balanced	right
House 968	Subset-Accuracy	.949	.952	.949
	F-measure	.755	.78	.741
House 992	Subset-Accuracy	.997	.997	.997
	F-measure	.775	.774	.78
House 938	Subset-Accuracy	.854	.854	.854
	F-measure	.51	.507	.501

When we consider a fixed sliding window, it is difficult to affirm which orientation works better from the results, as the difference is statistically insignificant. For that reason, the proposed model in the following results is based on the balanced orientation. Note that for the requirements of an on-line system (proposing real-time analysis), the left orientation should be preferred.

In future works a metric may be formulated using for example the duty cycle of the appliances under investigation to calculate the optimal window size. This period of time could replace the fixed size of the sliding window and lead to a better efficiency of the overall classification algorithms. The main problem being that the classification would be appliance-specific or well adapted for a group of appliances.

7.2 Energy variation between time steps

The energy variation between time steps is the one of the meta-feature presented in the Chapter 5, Section 5.2.2.3. This meta-feature is illustrated in Figure 5.3. As already discussed in that chapter, we have seen that meta-features play an important role in the proper classification of loads.

As an illustration, we propose in table 7.2 a sensitivity study on one of the basic meta-features, *energy variation between time steps*. The classification algorithm is run with this meta-features, or with only the knowledge of the energy reading from the power meter during the training session.

Table 7.2 – Impact on the results of using “Energy variation between time steps” or only “Energy” as features for the classification algorithm.

Test house	Measure	Energy variation between time steps	Energy
House 1	Subset Accuracy	.95	.95
	F-measure	.78	.77
House 2	Subset Accuracy	.997	.997
	F-measure	.77	.722
House 3	Subset Accuracy	.854	.821
	F-measure	.507	.36

It is seen that taking into account this feature works better than using only the exact energy values of the reference time in the sliding window. In the results, the subset accuracy are similar but the F-measure is higher. Note that a cross-validation evaluation is used to distinguish between features.

Similar results are obtained by considering the other meta-features one by one. This principle already has been illustrated in Chapter 4, Section 4.1.2.

7.3 Identifying the loads

The objective of this work is firstly to identify loads in buildings. Therefore, classification algorithms have to be compared on their ability to conduct such task. In this section, a comparison of some of the major classification algorithms is conducted on categories of appliances extracted from the houses of the Irise database. Here is a short summary of the chosen algorithms and their description. For further information on their principle and qualities, please refer to the Chapter 6.

BR1 : Binary relevance problem transformation using decision tree algorithm [Tsoumakas 2007].

BR2 : Binary relevance problem transformation using Svm algorithm [Tsoumakas 2007, Wu 2007].

LP1 : Label powerset as problem transformation and decision tree as base classifier [Tsoumakas 2007].

LP2 : Classifier chain algorithm using Svm as base classifier [Read 2011].

MLKNN : Multi-label K Nearest Neighbours with K=7 [Zhang 2007].

7.4 Load identification for defined categories of consumers

For a qualitative analysis of the database, the clustering proposed in chapter 3 is used for grouping. In this chapter, we have defined four clusters. In practice, we have also identified that two of these four clusters were close, one of them containing only three houses. We decided then to group them differently, based on their appliances configuration in addition than on their features dispositions. The houses from the three major clusters correspond to houses containing different appliance categories as follows :

Cat. 1 : Small number of distinct high energy appliances.

Cat. 2 : Small number of distinct high energy appliances with a few grouped appliances.

Cat. 3 : Many high energy appliances including grouped and repeated appliances.

These categories serve two purposes. The first one is to propose a relevant discrimination of the houses based on the facility to identify the state of all their high energy loads then to compare algorithms efficiency. Indeed, the grouping of loads has a significant impact on the algorithm efficiency. Therefore these categories are primary used here to compare the implemented algorithms. The second one is to gather potential information for the local grid manager. A category of houses reflect the potential flexibility level in term of aggregated power that could be obtained with appropriate tools.

The comparison of multi-label classification algorithms are presented in Section 7.4.1, 7.4.2 and 7.4.3 for the three categories of houses defined previously using F-measure as performance evaluation (refer to Section 6.5 for more information on the performance evaluation of the classification algorithms).

10 minutes and 1 hour sampling rates are presented together only for the first category of houses. A comparison of two multi-label learners with Hidden Markov Model is also proposed in Section 7.6.1, using Auc as performance evaluation. All the results discussed here are obtained using different readings of houses extracted from the 100 houses of the Irise database.

7.4.1 Considering the first category of houses

For the first category of houses, the results are shown using multi-label learners on different appliances for two houses grouped together. Two sampling times are used here in order to compare their impact on load identification.

The results proposed in table 7.3 indicate that the multi-label learners such as Lp and Ml-Knn (relying on appliance correlation) offer better performances on some of the appliances at a 10 minutes sampling rate. That is not the case for the other algorithms which do not consider appliances correlations. It can also be observed that generally, rule based algorithm such as Dtl as base learner provides better performances.

The results of the Table 7.3 can also be observed in the Figure 7.5 for better visualization.

The scores are much lower at a sampling rate of 1 hour and it is difficult to distinguish classifiers, except for the washing machine. Indeed, the correlation among appliances is weak and the learner is over-fitting during training, as it can be seen in the Figure 7.6.

TABLE 7.3 – Cat. 1 : Comparison of different multi-label algorithms for identifying the loads of a residential building.

Appliance	Sampling Rate	Algorithms				
		LP1	LP2	BR1	BR2	MLKNN
Washing Machine	10 min	75.72	72.50	79.32	74.49	71.98
	1 h	55.22	59.55	60.07	54.5	55.87
Microw. Oven	10 min	45.65	36.87	19.04	31.37	33.45
	1 h	29.94	11.40	21.13	5.10	2.12
Water Heater	10 min	96.66	95.99	97.35	92.02	97.40
	1 h	89.50	90.46	91.12	91.53	89.56
Electric Oven	10 min	66.79	55.82	73.65	46.81	60.53
	1 h	51.04	54.46	50.14	40.30	20.30
Clothes Drier	10 min	76.34	79.21	75.24	68.51	87.15
	1 h	58.16	60.65	62.21	60.11	69.50
Dish Washer	10 min	64.93	66.40	61.42	67.21	79.78
	1 h	36.86	36.50	36.72	34.66	32.63

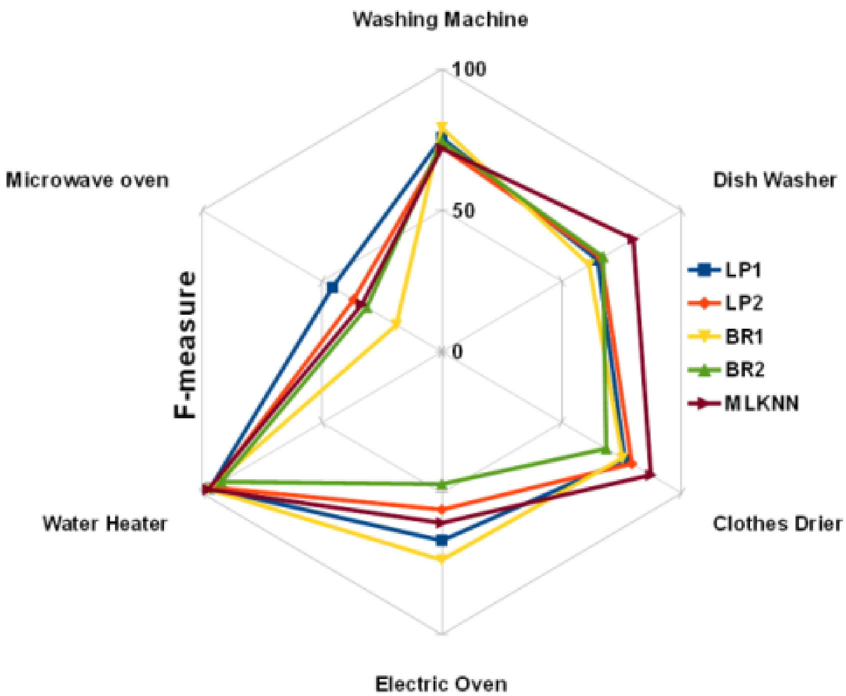


FIGURE 7.5 – Cat. 1 : Load identification performances for five different algorithms with a 10 minutes time sampling (algorithms acronyms in Section 7.3).

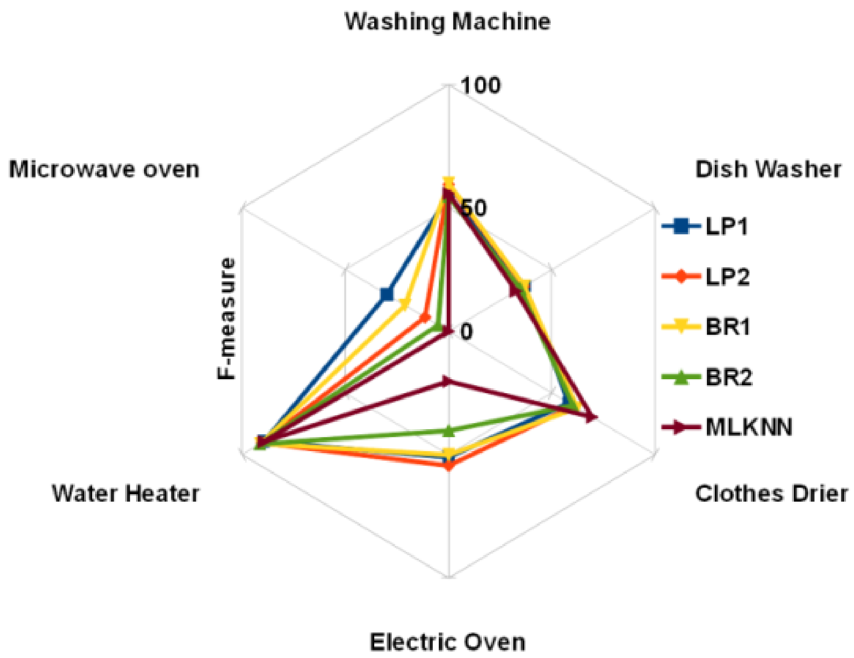


Figure 7.6 – Cat. 1 : Load identification performances for five different algorithms with 1 hour sampling rate (algorithms acronyms in Section 7.3).

Among all appliances water heater is identified with the highest scores and microwave oven with the lowest. This is due to the fact that even at a low sampling rate, the water heater keeps similar temporal usage patterns. On the contrary, the microwave oven presents high variations in both the duration of usage and the consumed energy which makes its identification difficult.

A microwave oven usage provides potential information for users if they want to reduce their energy consumption at a given time ; it is certainly not a controllable load. Actually, as a side result on load identification, only well identified appliances (washing machine, clothes drier, dish washer and water heater) present the possibility to become automatically controlled appliances. In fact, the water heater has already a possible distant grid control in France, the heating of the water operation depending on *term of use pricing*.

7.4.2 Considering the second category of houses

In this identification, the washing machine and the clothes drier are grouped together and considered as one appliance, along with other loads in the house. The sampling rate is of 10 minutes. The performance of the five algorithms is presented in table 7.4.

It is observed that when the numbers of high consuming appliances are small in the residence the performances are generally good. As in the previous section, the Lp algorithm and Ml-Knn present the best performances. Compared to the performances of algorithms if no load is grouped (Cat. 1), it can be seen that the appliances are identified with a better F-measure for all of the five algorithms. This category of household is particularly interesting considering identification algorithms and potential (distant) control for energy management and grid support through ancillary services.

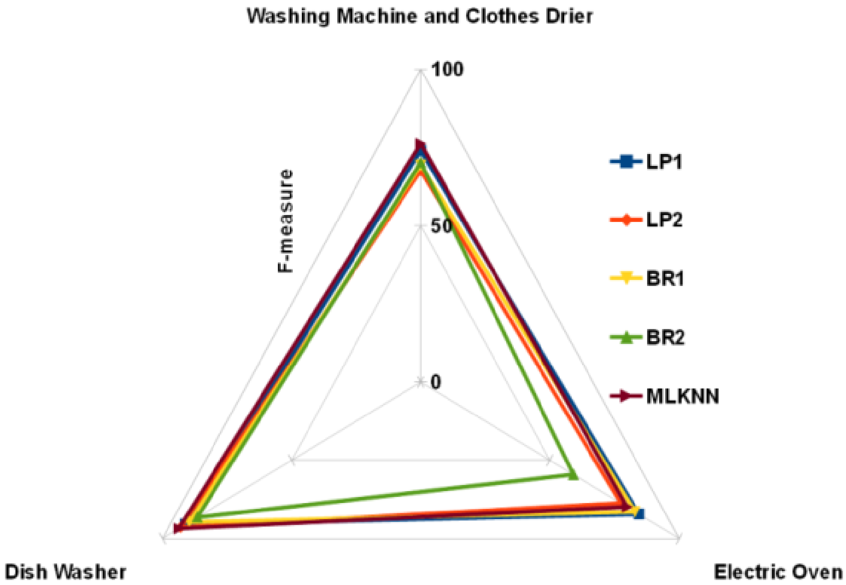


FIGURE 7.7 – Cat. 2 : Load identification performances considering some grouped appliances with a 10 minutes sampling time (acronyms in Section 7.3).

7.4.3 Considering the third category of houses

In this section, a high number of appliances including doubled appliances are considered. The results are shown in the Table 7.4 and the Figure 7.8, with a 10 minutes sampling time. The performance of the classifiers is drastically reduced as the number of appliances increases. Only the first water heater is identified with a sufficient accuracy, because its temporal features are well defined but it is not the case of the second water heater, which presents a lower number of uses.

At such low sampling rate, it is difficult for the classifier to learn the right temporal features of the appliances, if so many of them are used simultaneously and especially, if some of them are repeated. In that case, the short training phase is insufficient for a proper identification. Finally, it is observed that the single multi-label classifier (ML-KNN) presents lower performances than the transformation based multi-label classifiers using a single-label classifier at a core (for example BR1). This can be explained by the difficulty to find a specific correlation between so many appliances. Considering all labels at the same time are not of much interest in that particular case.

This work is based on voluntarily restrictive conditions. Two appliances state can change at the same time and the training period is short for the classification algorithms. Indeed, actual consumers would not accept a long period of monitoring and will use simultaneously their appliances.

Considering the results of all the three categories of houses, an algorithm capable of using appliances correlations is generally more suited than a one which is not. The Cat. 2 is the most promising one, based on the good identification capabilities associated with multi-label classification algorithms taking into account appliances correlations. On the contrary, when the number of appliances increases too much, then it is more efficient to suppress the possibility of taking into account appliances correlations because it interferes with the identification capabilities : the

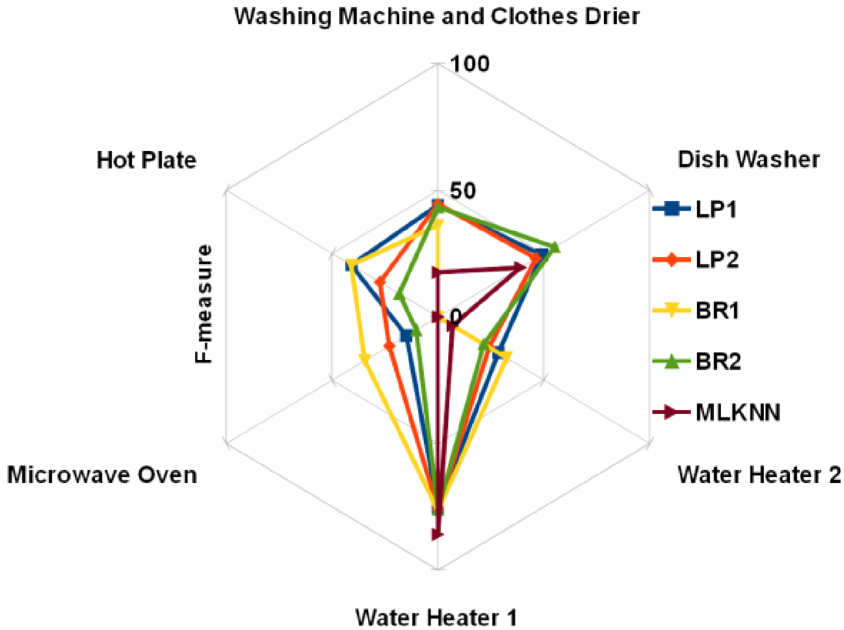


FIGURE 7.8 – Cat. 3 : Load identification performances considering some grouped and repeated appliances with a 10 minutes sampling time (algorithms acronyms in Section 7.3).

TABLE 7.4 – Cat. 2 & 3 : Comparison of different multi-label algorithms

Cat.	Appliance	Algorithms				
		LP1	LP2	BR1	BR2	MLKNN
2	Wash.Mach. Cloth.Drier	73.93	67.62	69.37	70.17	76.18
	Dish Washer	91.05	92.88	89.56	86.52	93.65
	Electric Oven	84.46	77.44	82.92	59.06	79.81
3	Wash.Mach. Cloth.Drier	44.14	44.55	36.13	43.27	17.56
	Hot Plates	40.96	27.43	40.96	18.31	2.14
	Microw. Oven	14.93	23.61	34.56	10.29	1.32
	Water Heat. 1	75.17	76.33	76.56	76.16	85.92
	Water Heat. 2	28.49	24.11	32.43	21.58	6.90
	Dish Washer	48.93	46.23	42.16	55.10	38.93

classification algorithm is trying to find relations where there is none.

7.5 Sensitivity to the sampling rate

One of the key points of the work is the choice of keeping the sampling rate at 10 minutes and to test further lower ones (30 minutes or 1 hour). Indeed, depending on the sampling time, the computation workload and the need for big-data algorithm varies a lot. Testing the sensitivity of the classification algorithm to these sampling rates is of much interest considering the possible sampling rates of future smart meters.

In Fig. 7.9 the subset accuracy (refer to Section 6.5) for four houses is shown with a 10 minutes, 30 minutes and 1 hour sampling rates. The score logically increases with the sampling rate, but the variation is depending on the considered house.

A confirmation is obtained through the comparison of the results presented in Fig. 7.5 and 7.6. For high energy consuming appliances, increasing the sampling rate from 30 minutes to 10 minutes is a benefit regarding the classifier performance.

The performances of the classifiers vary from one house to another, depending on the temporal behavior of the inhabitants (more regular or not). This conclusion has been validated on all of the 100 houses of the IRISE database.

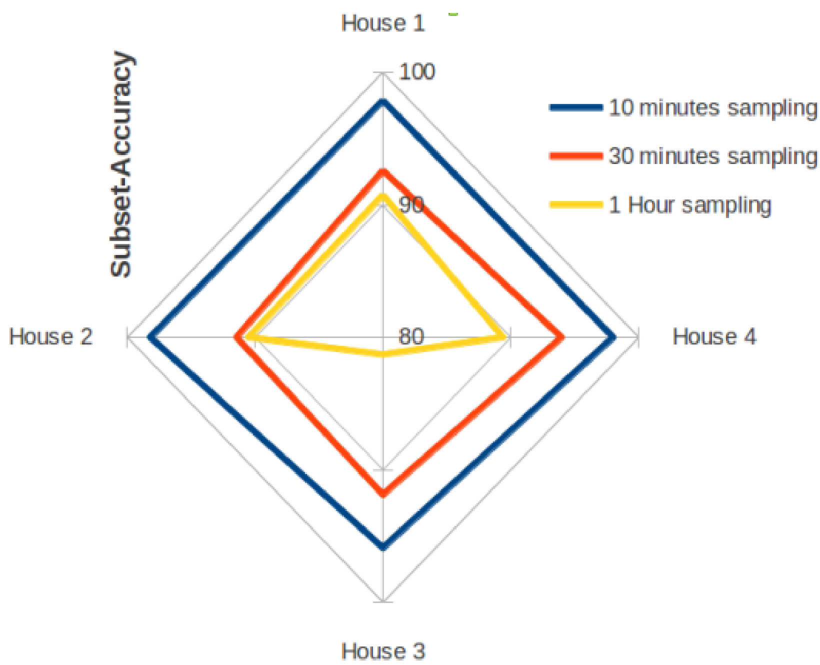


FIGURE 7.9 – Load identification subset accuracy for 3 sampling rates and 4 houses.

7.6 Comparison with standard non-intrusive load identification methods

7.6.1 Hidden Markov Model

In table 7.5, a comparison is shown between two multi-label learners and a standard non-intrusive load monitoring algorithm (HMM). The number of states of the HMM is experimentally set to 2 and the input is the time series subsequence. The results are shown for four houses and tend to hold generally for the whole dataset. It indicates that at low sampling rates the multi label learners using meta-features perform considerably better than HMM. LP2 which considers appliances correlations performs better than BR2 and it is seen that SVM as a base learner performs better (considering AUC measures).

TABLE 7.5 – Comparison of Multi-label learners and Hidden Markov Model with AUC measures

Residence	Algorithms		
	BR2	LP2	HMM
House 1	97.90	99.24	80.20
House 2	95.57	98.42	72.27
House 3	89.76	97.51	63.55
House 4	95.27	98.96	84.67

7.6.2 KNN Distance based measures

One representative house from each of the clusters has been selected to analyze the impact of different type of algorithms. A comparison is made for the different algorithms for each house. The best performing algorithm is selected and results for each individual appliance is presented in table 7.6. The distance metrics were discussed in details in section 6.2.

TABLE 7.6 – Comparison of distance based metric and Hidden Markov Model with AUC measures

Residence	K-Nearest Neighbours Distance Measures			
	DTW	TC	Euclidean	HMM
House 1	.91	.50	.83	80.20
House 2	.85	.78	.87	72.27
House 3	.82	.70	.78	63.55
House 4	.81	.80	.86	84.67

The results in 7.6 shows that the distance based time series metrics are performing better than HMM. This could be attributed to many factors, the primary among them is the fact the DTW is based on similarity between two temporal sequences which may vary in time or speed and TC can take into account the similarities between two behavior based metrics.

DTW as a measure is performing best for two houses and the default euclidean distance for the other two. This indicates the fact that in this domain appliance signature may be more influenced by variation in time and speed than by behavior. Looking internally into the results, temporal correlation seems to wrongly classify appliance ON state of one appliance and pass it to another, as the two signatures have a similar behavior but differ by magnitude.

HMM results are the lowest among the tested algorithms. The low performances of HMM can be attributed to the possibility that many of the appliances in the database do not always show obvious indication of the times when they are switched ON.

There are different appliance load patterns. A similarity-based or statistically-based learning scheme should be able to correctly classify these groups, given sufficient training examples from both sets (ON and OFF states). A similarity-based learner would be able to classify the non-obvious examples separately, using whatever similarities did exist between the instances of each class. A statistically-based system cannot do this. In that case, HMM will attempt to build a single model that accurately describes all of the training examples for a particular class. The examples from the appliances signatures that do not show any recognizable pattern contribute with very little information about the class, but the model will still attempt to account for them. This behavior leads to both models trying to classify a group of quite similar examples. Consequently, the models will perform poorly when an instance from this group is encountered during testing. We used a discrete markov model for supervised learning but there exists other ways of modelling using Hmm which needs to be further explored.

In table 7.7 the F-measure (section 6.5.1) for different appliances present in different houses are shown. The results are proposed taking each time the best performing Knn algorithm (with different metrics) for each appliance in each house.

TABLE 7.7 – Appliance level performance with F-measures (N.A. is used for "Non Applicable").

	Houses			
Appliance	House1	House2	House3	House4
Water Heater	.91	.94	.91	N.A
Washing Machine	.30	.82	N.A	N.A
Dish Washer	.43	N.A	.60	.86
Electric Oven	N.A	.56	N.A	.36
Microwave Oven	N.A	.18	N.A	.01
Clothes Drier	N.A	N.A	.39	N.A
Washing Machine+Clothes Drier	N.A	N.A	N.A	.58
Electric Cooker	.64	N.A	N.A	N.A

Table 7.7 gives an insight of the high consuming appliances present in the house as well as their predictability. In can be seen from the results that water heater is classified the best, followed by dish washer, electric cooker and electric oven. These appliances have a significant amount of consumption even at a low sampling rate so they can be classified better. Devices such as washing machine and washing machine + clothes drier show an intermediate performance. For some houses then can be classified better than others, this can be attributed to the different

possible operational modes of the device. Just a mere "binarisation" of the states to On/Off does not work well for all the cases. Microwave oven is the less identified appliance among the devices, as it can be used in multiple modes and different duration so its identification task is very difficult at this sampling rate.

7.7 Discussion

The results indicate that for high power consuming appliances the proposed methodology gives reasonably good results considering the large span of time it is tested. In a particular period of training (more than a few months), the shorter the period of testing is, the more accurate are the results.

The predictability of the states (i.e. their identification) also varies among the appliances. Some appliances such as water heater are easier to disambiguate than other such as washing machine or microwave oven.

This fact indicates the necessity to group these appliances into more abstract groups. So *Washing* can be a group for washing machine and clothes drier, similarly *Cooking* can regroup the microwave oven and the hot plates. It depends on the appliances used by the domestic user for such groupings, if high consuming appliances are small in number and only from one group then grouping might not be necessary.

It also needs to be pointed out that other appliances in the house are neglected at this sampling rate : low consuming appliances. Maybe at lower sampling rate these appliances may also be predicted. The approach is based on the assertion that instances seen in the future can be classified based on what is seen in the past, which has its own limitations.

The proposed Nialm technique is suitable for scenarios where multiple appliances start at the same time (similar and dissimilar). The method used in this work is non-event based and uses a multi-label classification approach thereby developing a separate category when two appliances are in the On state. When the washing machine and the water heater are working during the same period of time, a new combined binary class label will be generated representing the washing machine and the water heater On states and will be compared with similar instances encountered during training. This holds true also for two similar appliances, for example if two appliances with two possible states (On and Off) are represented as binary "1" and "0" respectively. Four new classes will be generated represented as "00", "01", "11", "10" considering all possible state combinations.

Most of the common algorithms in the multi-label classification have been tried and in general the label powerset (Lp) as problem transformation and decision tree (Dtl) as the base algorithm works better. This may be attributed to the fact that the number of labels is not very large. The computation cost of the algorithm is also quite low.

Conclusion

Household's energy management is an important discipline considering the impact of energy efficiency on different levels from global policies to particular behaviors. The work proposes a multi-label learning method that takes into account appliances correlations based on novel meta-features in order to identify loads without extensive monitoring of inhabitants during the training phase and without any monitoring thereafter, except for the active power meter measurements, all this at a low sampling rate. Inhabitants may then monitor their energy consumption for a short period of time and subsequently get an energy management plan for the rest of the year. For grid managers, a good identification will lead to better possibilities of flexibility assessment and requests (through distant shut down, load shade or shift) and also a better global behavior prediction, i.e. without intervention into resident's private life.

The results are computed using 10 minutes and 1 hour sampling rate on the Irise database (including 100 houses monitored over one year) using a range of multi-label learning algorithm. The choice of this sampling rate is done in order to avoid privacy issues, to stay with realistic order of magnitude considering the first generation of smart meters and to decrease the needs for big data.

The results indicate that consideration of temporal knowledge leads to an increased capability of non-intrusive disambiguation of the aggregated load. The use of multi-label learners also exhibits that there are appliance correlations.

The presented algorithms are well suited for load identification, considering particular hypothesis (like appliance grouping) that allows defining categories of houses, for example, small number of appliances, different high energy appliances, repeated appliances, etc. The learned models are also of interest regarding the future work, which will concern energy management applied to smart buildings and behavior prediction.

Quatrième partie

Appliance Usage Prediction

Chapitre 8

Methodology of Appliance Usage Prediction

8.1 Load forecasting algorithms **89**
8.2 Problem statement **94**
8.3 Overall proposed architecture **96**
8.3.1 The oracle 96
8.3.2 The predictor 97
8.3.3 Data Selector 97
8.4 The oracle : for knowledge extraction **98**
8.4.1 An expert system that generates knowledge : the oracle 99
8.4.2 The oracle output 101
8.5 The predictor : for classification **101**

The objective of this work is to statistically predict the user energetic service request for the next 24 hours using an enrich learning algorithms with expert knowledge. We sample the time space into 24 hour and wish to predict the user appliances usages requirements for a particular hour. At each point of time the prediction system will predict the following 24 hours and then shift to the next hour and predict the following 24 hours one again, etc.

The novelty of our approach lies in the proposed general model which is still lacking in the domain of appliances usages prediction for home automation systems. It is a difficult problem because prediction at appliance level does not benefit from any profusion. It is difficult to tackle these prediction problems with usual prediction approaches. Nevertheless, the algorithms of prediction are tested in the Irise database and the results indicate that they are well suited for this application.

8.1 Load forecasting algorithms

The problem of forecasting has been a subject of research for a considerable number of years. The workability of a technique depends on its simplicity and comprehensibility (i.e. the meaning) of the model being used. Forecasting has been used in a number of domains, at first, one needs to look into the domain of *load forecasting* at the grid level and review the approaches used. In this chapter, a brief summary of load forecasting models is provided and sub-subsequently the model used for residential load is discussed in details.

Load forecasts can be divided into three categories :

Short term forecasts : From one hour to one week.

Medium forecasts : From a week to a year.

Long-term forecasts : Longer than a year.

The problem of appliance usage prediction has similarity with the short-term load forecasting (STLF). We list here some of the common approaches used in the later. Though STLF uses regressive approaches whereas our approach is based on classification, the strategies used in the domain of energy load prediction will guide us in our choice of input to the predictor.

In the following paragraphs, we look into the approaches already used in load prediction [Feinberg 2005]. The approaches are not independent of each other but rather complementary. An important aspect of forecasting algorithms is that the simpler models are more popular among utilities, of course without being too trivial. The features of a forecasting algorithm depend on the business needs it is implemented for. The challenge is not only to be technical but analytical in the approach to build a proper model of forecasting tool.

Similar-day approach : This approach is based on searching historical data for days within one, two, or three years with similar characteristics to the day to forecast. Similar characteristics include weather, day of the week, and the date for example. In the case of a load forecasting, the load of the similar day is taken as a forecast [Feinberg 2005].

This process is simple to comprehend but in relevant cases it can outperformed more complex mathematical approaches. It also must be mentioned that, though this method is simple, it is not trivial and requires a good understanding of the domain of interest.

Regressive methods : For electric load forecasting, regression methods are usually used to model the relationship of loads consumption and other factors such as weather, day type, and customer class. [Engle 1992] presented several regression models for the next day peak forecasting. For appliance prediction researches indicate that a classification based on appliance categories would be more applicable [Hawarah 2010].

Indeed, the energy consumption values are highly random to be predicted correctly in a regressive approach rather than a class based approach to determine the appliance state followed by the assignment of the energy value.

Time series : Time series methods are based on the assumption that the data have an internal structure, such as autocorrelation, trend, or seasonal variation. In particular ARMA, ARIMA and ARIMAX are the most often used classical time series methods [Tran 2012]. The idea of the time series approach is based on the understanding that a load pattern is nothing more than a time series signal with seasonal, weekly, and daily periodicities.

Generally, techniques in time series approach work well unless there is an abrupt change in the environmental or sociological variables which are believed to affect load pattern. Our work takes into account the time series approaches using a Neural Network predictor [Tran 2013] but it must be emphasized that, unlike in time series based energy forecasting our approach is classification based where the energy-values are discrete in time. Neural networks are also addressed in a following paragraph.

Expert systems : Expert systems incorporate rules and procedures used by human experts in the field of interest into software. From that knowledge, these softwares are able to automatically make forecasts without human assistance. Knowledge-based expert system for the short-term load forecasting have already been successfully deployed in the world, for example for the Taiwan power system [Ho 1990]. In this example, operators knowledge and the hourly observations of system loads along with weather parameters were taken into consideration.

Our model also proposes a general model which takes the expert knowledge into account. The detailed approach regarding how to formalize expert knowledge is discussed in chapter 9.

Artificial neural network, ANN : Artificial neural networks have also been applied in the domain of energy load forecasting. The studies conducted in [Hippert 2001, Bakirtzis 1996, Park 1991, Khotanzad 1998] give an adequate idea of the architecture and parameters that are most commonly used for energy load forecasting.

Neural networks are essentially non-linear circuits that have demonstrated the capability to do non-linear curve fitting. The outputs of an artificial neural network are some linear or non-linear mathematical function representing its inputs. The inputs may be also the output of other networks elements. In that configuration, they are arranged in a relatively small number of connected layers of elements between networks inputs and outputs.

As an illustration, the Figure 8.1 proposes the input configuration of an implementation of neural networks in Short term load forecasting.

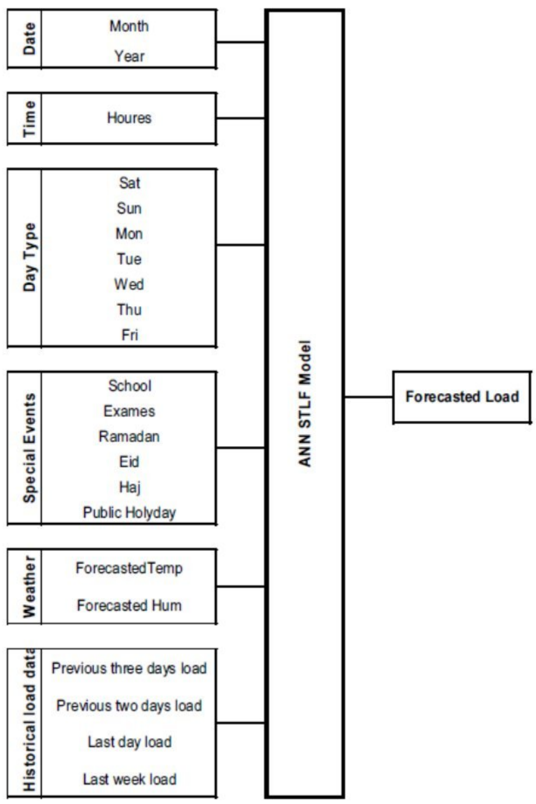


FIGURE 8.1 – Short term load forecasting predictor input configuration

The neural network architecture used in the prediction is a multi-layer perceptron. The ANN model is feed forward back-propagation learner that maps a set of input to a output model. For load forecasting the commonly used learning methods can be either Hopfield, Back propagation or Boltzmann machine. In our work, the Back propagation learning [Hecht-Nielsen 1989] was used. The final choice of all the parameters is shown in the Table 8.1.

TABLE 8.1 – Artificial Neural Network parameters chosen in these works.

Network Type	MLP
Activation function (hidden unit)	Tanh
Activation function (output unit)	Softmax
No of hidden neurons	no of input/2
Error Function	Cross entropy
Training Algorithm	BGFS
Learning Rate	0.1

Support vector machines : Support vector machines perform a non-linear mapping (by kernel functions) of the data into a higher dimensional feature space. Then the algorithm uses simple linear functions to create linear decision boundaries in the new space. SVM model can be used to predict daily load demand, for example for the next month [Chen 2001].

The problem of choosing an architecture for a neural network is replaced here by the problem of choosing a suitable kernel function for SVM. The detailed implementation of the support vector machines is discussed in chapter 5.

Decision table : Given a training sample containing labeled instances, an induction algorithm builds a hypothesis in some representation. The representation investigated here is a decision table [Kohavi 1995] with a default rule mapping to the majority class, which is abbreviated as DTM. A DTM has two components :

- A schema, which is a set of attributes.
- A body, which is a set of rules (labeled instances). Each rule consists of a value for each of the attributes in the schema and a value for the label.

Given an unlabeled instance I, the label assigned to an instance by a DTM classifier is computed as follows. Let L be the set of rules in the DTM exactly matching the given instance I, where only the attributes in the schema are required to match and all other attributes are ignored. If $L = \phi$; return the majority class in the DTM otherwise, return the majority class in L.

The most important part of the decision table are the rules. Some of the rules used by decision table classification for the oven consumption are given in figure 8.2.

In this particular situation, the most important attributes considered when building the rules are : the hour and the consumption at hours $H - 11$, $H - 2$ and H. Based on this rules, the unlabeled instance "$consH + 1$" will be classified.

Rules:

Hour	ConsH-11	ConsH-2	ConsH	ConsH+1
23	1	1	1	0
17	0	1	1	0
10	0	1	1	0
9	0	1	1	0
14	0	1	1	1
20	0	1	1	0
18	0	1	1	1
16	0	1	1	0
19	0	1	1	0
22	0	1	1	0
12	0	1	1	0
15	0	1	1	1
11	0	1	1	1
13	0	1	1	0
21	0	1	1	0
0	0	1	1	0
23	0	1	1	0
7	1	0	1	0

FIGURE 8.2 – Decision table rules used for oven usage prediction

Bayes Net : A Bayesian Network (BN) is a graphical model for probabilistic relationships among a set of variables [Pearl 1986]. The BN algorithm models causal relationships. They are represented as directed acyclic graphs, where each node represents a different random variable. A directed edge from the node X (the *cause node*) to the node Y (the *effect node*) indicates that X has a direct influence on Y. This influence is quantified by the conditional probability $P(Y|X)$, stored at node Y.

The nodes in a network can be of two types : *evidence node* when its value is observed, and *query node* when its value has to be predicted. A Conditional Probability Table (CPT) is assigned to each node in the network. Such probabilities may be set by an expert or using a registered data. BN are based on the conditional independence. Each node is conditionally independent of its non-descendants given its parents. When a node has no parent, its CPT specifies the prior probability.

There are two types of learning :

1. *the structure learning* in which the best graph representing the problem is searched
2. *the parametric learning* in which the structure of the network is known and the conditional probability is estimated at each node.

Basically, the classification task consist of classifying a variable called the class variable given a set of variables called attribute variables [Witten 2011].

A simple example of Bayesian network classifier is presented in figure 8.3. The last 4 hours historical data concerning the energy usage of the appliance electric oven in one of the houses of the IRISE database is considered. The class variable is the energy usage for hour H and the attribute variables are hour, day, consumption at hour $H-4$ (named $Cons_{H-4}$), etc. The CPT is computed for each node of the network, the figure shows the one for the energy usage in the hour $H-3$.

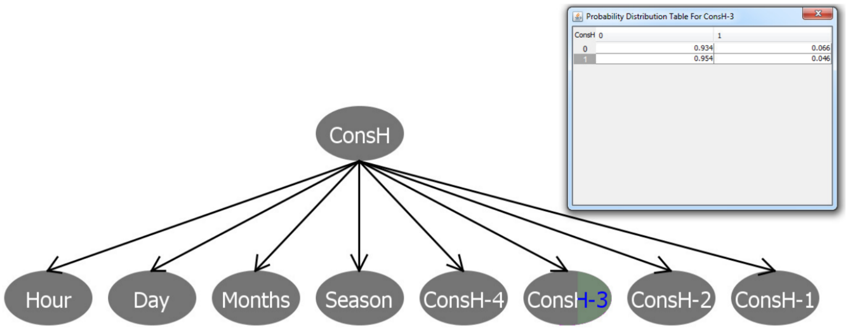


FIGURE 8.3 – Bayes network for the electric oven in the house N^o 997 of the IRISE database.

8.2 Problem statement

The energy management system goal is to advise the inhabitants to consume in a certain energy efficiency spirit or in order to reduce the energetic bill [Basu 2013b].

As already discussed in the Chapter 2, the EMS considered in this work is a software which architecture has three layer :

The anticipative layer is responsible for the scheduling end-user, intermediate and support services, taking into account predicted events and costs in order to avoid as much as possible the use of the reactive layer. This layer has slow dynamics and includes predictive models with learning mechanisms and is able to compute energy consumption plans with several hours in advance.

The reactive layer objective is to manage adjustments of energy assignment in order to follow up the plan computed by the anticipative layer despite unpredicted perturbations. Actions of the reactive layer may be to delay the starting time of an appliance, to interrupt or even to stop an appliance, etc.

The local layer is composed of devices together with their existing local control systems generally embedded into appliances by manufacturers. It adjusts device controls in order to reach a given predefine set points.

An important part in computing the energy plan for the next day is to consider the energy consumption prediction. The objective of this work is to propose a *learning system* able to help the *Home automation system* to compute the energy plan taking into account the user requests. The focus is set on the prediction of the appliances usage in buildings and their energy consumption. The predictor determines whether the service will be consuming energy or not. The goal is to predict, in the next hour, the start/no start of the appliance.

It is easier to predict the total energy consumption of a house than to predict energy consumption for a single appliance. As an illustration, figure 8.4 shows that by using a weekly database, it is possible to predict the energy consumption of an entire house.

This figure represents the distance between the consumptions of two weeks over a whole year. One week on the X-axis and another one on the Y-axis. The distance is computed considering

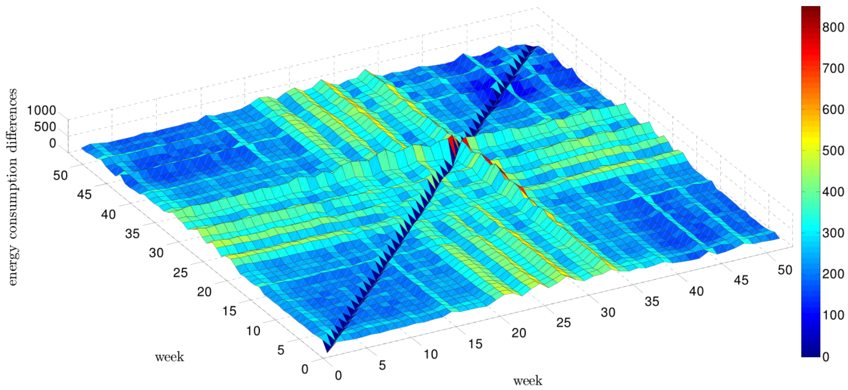


FIGURE 8.4 – Correlation between weeks for the total energy consumption of a house over a year.

the energy consumption at each hour of the week, using the euclidean distance defined as :

$$d_E(X,Y) = \sqrt{\sum_{i=1}^{24*7}(x_i - y_i)^2} \tag{8.1}$$

where X and Y are vectors representing the energy consumption components. x_i and y_i are the energy consumption in Watt.hour for each hour i. These variables are independent one to each other since the hourly consumption is random, depending mainly on inhabitants behavior.

It becomes more complicated when only one service is predicted as seen in figure 8.5. The picture concerns the washing machine in the same house and represents the distance between the energy consumed each two successive weeks.

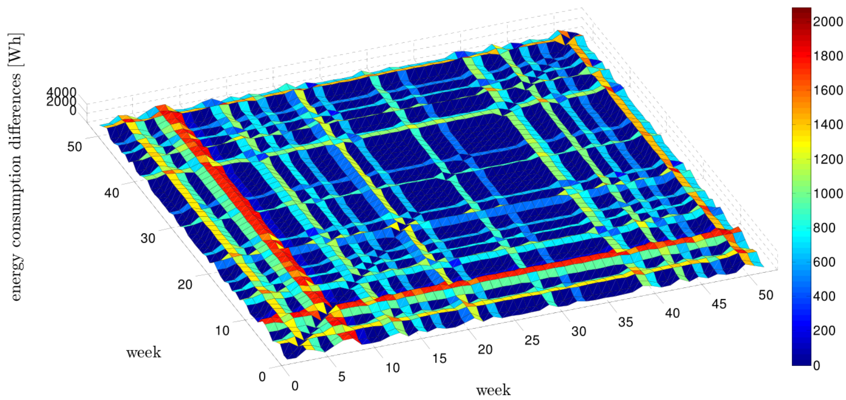


FIGURE 8.5 – Correlation between weeks for only the energy consumption of a washing machine in one house over a year.

By comparing the Figures 8.4 and 8.5 we can conclude on the difficulties encountered by the prediction algorithms once we decide to move inside the buildings instead of only working on global consumption forecasts.

From a prediction point of view, an appliance in a house is a service to be used by inhabitants. For each service and for a chosen time period, the profile of prediction consists of :

- the appliance is used during that hour
- the average duration every hour,
- the average consumed energy every hour,

The duration or the energy for a service is assumed to be fixed for an hour. The primary task of this research is to predict if an appliance is used during a hour. It can also be useful to calculate the probability that an appliance starts at each hour over all the year without taking into account any time period. This kind of information can be used to briefly depict the profiles and then identify the most appropriate profile to a given user.

The results presented in figure 8.6 shows the probability that the electric oven starts in a house. At 11am it is of 22% and at 6pm of 35%. These values are calculated over one year of observation.

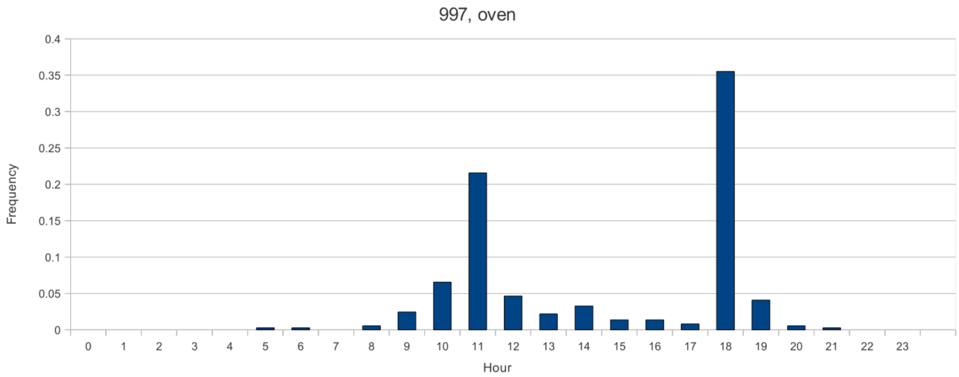


Figure 8.6 – Statistical probabilistic representation of the consumption of an electric oven (house N^o 997 of the Irise database).

8.3 Overall proposed architecture

The model developed in the work consists of an improved learning algorithm which proposes a general way to take expert knowledge into account. This model divides the processes into modules. Each module has its own purpose. Subsequently, a model with a simplified, modular architecture is proposed as seen in the Figure 8.7. Finally, a real time multi-label approach is tested using the above model (refer to the Chapter 10). In the following sections each of the processing modules is discussed.

In this work, *information* refers to meaningful data related to a current situation. It is modeled as data and variables. *Knowledge* refers to the rules and relationships between information considered as valid by a person or a group of people. It is modeled by logical propositions. Raw data contains energy consumption for an appliance and contextual information, like the time, the date, the weather, etc.

8.3.1 The oracle

A simple model for prediction could be one which makes the forecast using a classification method based only on consumption information (i.e. just raw data). A model without any other

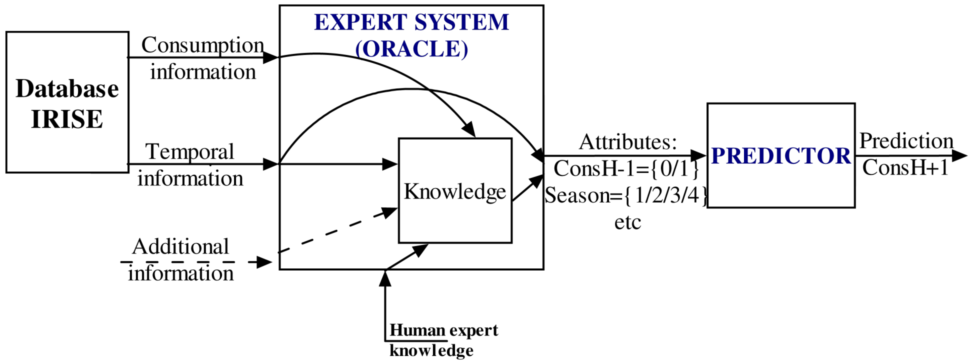


FIGURE 8.7 – Prediction architecture

expert knowledge is shown in the Figure 8.8.

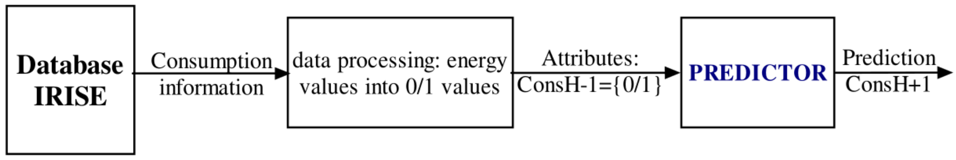


FIGURE 8.8 – Simple prediction architecture with no expert knowledge addition.

Usually, an “expertise” process is added to such predictive system, in order to simulate the intervention of a human expert, adding particular knowledge to the prediction. This we call as *oracle*.

The oracle generates additional information, according to expert knowledge. It receives the raw data from the database giving the consumption at a particular hour and in addition the date, time and weather information for that hour. The oracle proposes attributes based on knowledge and then gives the necessary function which represents the data in an interpretable form to the Prediction system.

8.3.2 The predictor

The Predictor consists of classifiers commonly used in machine learning such as the Neural networks, Bayes network, Decison tree, etc. The predictor receives the attributes based on domain knowledge proposed by the oracle. At this stage, the prediction task is similar to other machine learning classification task. The details about the classifiers and the measures used are detailed in the two previous chapters. The Figure 8.9 propose a first example of prediction conducted on the lighting in a house by using a simple forecaster. It can be observed that the prediction of real values (regressive approach) did not work well for appliance prediction.

8.3.3 Data Selector

We define a data selector as a non-temporal matrix processor which stores, selects and structures the data for the predictor. The produced matrix is the input to the predictor. The data selector may choose the whole or the subset of the output from the oracle and is controlled by the predictor controller.

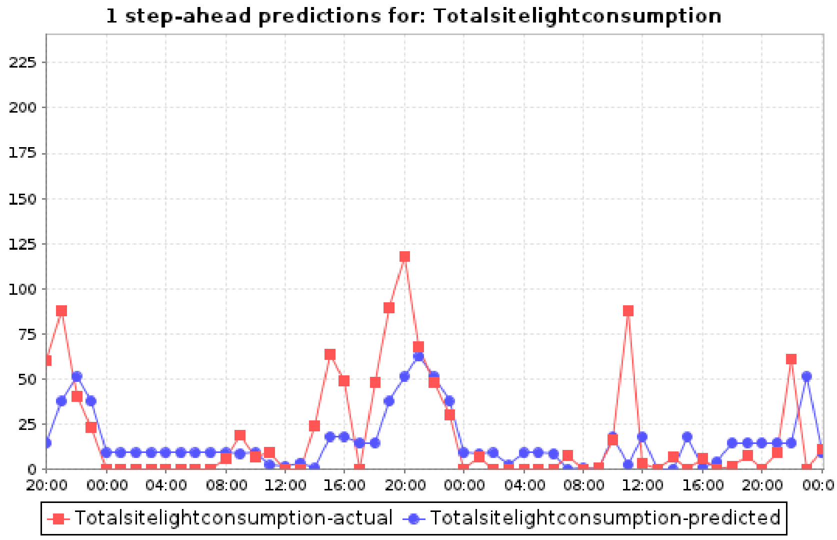


Figure 8.9 – One day ahead prediction for the complete lighting of a house.

It should be noted that all the knowledge proposed by the oracle might not be useful for a particular appliance in a house. There are different possible ways of structuring the knowledge proposed by the oracle. All the outputs of the oracle are stored in the data selector, but only those which are validated by the predictor are selected for the overall prediction. This is "decided" by the predictor controller.

In this work, the chosen way of structuring the output is done by taking all of the knowledge proposed by the oracle as a single unit. There are other possible ways of structuring that output, which will be looked in future works.

8.4 The oracle : for knowledge extraction

Knowledge extraction is a novel attribute construction technique that can be applied whenever there is some kind of underlying substructure to the training instances and there is some way to extract these substructures. In temporal domains, these substructures take the form of sub-events, like periodicity in data. Through the oracle, these substructures become synthetic features, which are then fed to a propositional learner. The output of the learner can then be converted back to a human-readable form that is described in terms of decision tree for example.

In the proposed model, the problem is reduced to the knowledge of an appliance being On or Off in a certain time interval. This simplification of the problem is done to reduce the uncertainty associated with the prediction of exact energy consumption values. Though, it must be mentioned that problem can be increased from binary class to multi-classes, to predict energy levels instead of only the loads states.

As shown in the Figure 8.9, the outcome of a prediction of the energy values by using the *weka* forecaster plug-in [Weka 2014] exemplifies the importance of reducing the problem into a simpler form, in our case, the simple state of the loads.

8.4.1 An expert system that generates knowledge : the oracle

We define the oracle knowledge as the statements leading to entities (or factors) that need to be taken into account for prediction. The oracle receives the raw data from the database giving the consumption at an particular hour plus for example the date, the time and the weather information at that hour.

The oracle integrates this knowledge and then transmit the necessary function which represents the data in an interpretable form to the prediction system. The knowledge which is relevant for a particular appliance in a house might not be relevant for another house using the same appliance. So all knowledge proposed by the oracle has to be validated and knowledge which doesn't increase or reduces the accuracy of prediction for a particular appliance will be rejected. The validation and structure of the output of the oracle is done in the subsequent processing module as seen in the Figure 8.7.

Some examples of statements that could be used by an oracle for a functional representation (being meaningful to the appliance usage prediction in our case) are listed below :

- The immediate past history of the consumption.
- The hour of the day.
- The day of the week.
- The season of the year.

The oracle knowledge is not restricted only to the implemented functions. For example, the lighting usage could be an indicator of occupancy and is related to all the other appliances. Additional information from the smart phone calendar could be integrated. In this application, the focus in on some of the cases, a practical model will need to model all the possible information. In the following paragraphs, we have a look at each of the knowledge proposed to the Oracle and their structural representation.

Past consumption history, same day, previous hours The past sequence of energy consumption prior to an event is meaningful in appliance usage prediction and is fed to the oracle. Mathematically, it is formalized by the following predicate function :

Input : $\{E(H-1), E(H-2), \cdots, E(H-n)\}$ where, E is the energy consumption, n is the size of the past time history and H the current hour.

Output : $\{0,1\}^n$

Here the output is a threshold binary vector of size n which signify if there is consumption at the hours prior to the event on not. As an illustration, consider an appliance used every three hours. To take into account this knowledge, we need to know the consumptions prior to that event. In this case, we will find out that for an event where the appliance was switched On, it was also switched On three hour before that event.

Hour of the day Through this knowledge, the time space is discretized into 24 hour slots and the "actual time" where the event occurs assumes priority.

Input : Current Hour in the day, $H \in \{(0-1), (1-2), ..., (23-24)\}$

Output : $\{0,1\}^{24}$

There is both the possibility of using orthogonal representation or using real values. For example, instead of using the real numeric value 6 O'Clock, we could use the orthogonal vector "$\{0,0,0,0,0,1,0,...,0\}$" to represent the same hour. As an illustration, a full day would represented in orthogonal vectors as proposed in the Table 8.2 :

TABLE 8.2 – Orthogonal representation of the hour of the day.

Hour 0	—	$\{1,0,0,0,0,...,0\}$
Hour 1	—	$\{0,1,0,0,0,...,0\}$
$\vdots$		$\vdots$
Hour 23	—	$\{0,0,0,0,0,...,1\}$

Taking into account the hour of the day will likely lead the predictor to associate the usage of a load with the behavior of the inhabitants. For example, appliances are always used one we wake up, like the coffee machine, etc.

Day of the week Similar to the way we take into consideration the hour of the day, the whole week can be discretized into 7 days.

Input : Current day of the week, $D \in \{\text{Sunday}, \text{Monday}, \cdots, \text{Saturday}\}$

Output : $\{0,1\}^{7}$

Instead of representing this with a numeric value, we use the also an orthogonal representation. So for a full week, the representation will be as presented in Table 8.3.

TABLE 8.3 – Orthogonal representation of the day of the week.

Sunday	—	$\{1,0,0,0,0,0,0\}$
Monday	—	$\{0,1,0,0,0,0,0\}$
Tuesday	—	$\{0,0,1,0,0,0,0\}$
Wednesday	—	$\{0,0,0,1,0,0,0\}$
Thursday	—	$\{0,0,0,0,1,0,0\}$
Friday	—	$\{0,0,0,0,0,1,0\}$
Saturday	—	$\{0,0,0,0,0,0,1\}$

This knowledge will lead to a potential prediction of a different behavior between weekdays and weekends.

Season of the year Similarly to the prior sub-sections, there are appliances in houses which show distinctive behavior depending on the season of the year.

Input : Current season of the year, $S \in \{\text{Spring}, \text{Summer}, \text{Autumn}, \text{Winter}\}$

Output : $\{0,1\}^{4}$

As for the prior knowledge, we choose an orthogonal representation over a numeric one. So the season oracle output will be presented as in Table 8.4

TABLE 8.4 – Orthogonal representation of the season of the year.

Spring	—	$\{1,0,0,0\}$
Summer	—	$\{0,1,0,0\}$
Autumn	—	$\{0,0,1,0\}$
Winter	—	$\{0,0,0,1\}$

This knowledge will be taken into account if we sample many years in terms of seasons. This will lead for example to a potential prediction of the electric heater used primarily in the winter.

Past consumption history, previous days, same hour The objective, is to assess if there is consumption or not in the previous days at the same hour. We typically chose to analyze the seven days before an event. The output is a vector of threshold binary values.

Input : $\{E(H-24), E(H-48), \cdots, E(H-m)\}$ where, E is the energy consumption, m is typically taken as 168 and H the current hour of the day.

Output : $\{0,1\}^7$

For example, this knowledge will lead to the potential prediction of the fact that people will eat every day at a particular hour, using their electric oven every day at the same hour.

8.4.2 The oracle output

The overall oracle output after the representation of the knowledge proposed as input is shown in table 8.5 where each row represent the proposed knowledge at a particular time by using feature selection.

The table is obtained by feature selection by the oracle (the knowledge proposed in the previous paragraphs are added or discarded depending on the feature selection).

8.5 The predictor : for classification

This module consists of the classifiers commonly used in machine learning. The classifier gets its input from the data selector. A classification learner approximates a function, mapping a vector into labels by looking at input-output examples of this function. The features x_i and the target class Y come in record of the form :

$$(x, Y) = (x_1, x_2, x_3, \ldots, x_n, Y) \tag{8.2}$$

The used predictors come in various categories of classification algorithms. These algorithms have been presented in the two previous chapters.

TABLE 8.5 – The overall oracle output.

Knowledge			Value
Past consumption history same day, n previous hours	$E(H-1)$	—	0 or 1
	$E(H-2)$	—	0 or 1
	$\vdots$	—	$\vdots$
	$E(H-n)$	—	0 or 1
Past consumption history m previous days, same hour	$E(H-24)$	—	0 or 1
	$E(H-48)$	—	0 or 1
	$\vdots$	—	$\vdots$
	$E(H-m)$	—	0 or 1
Hour of the day	$H(0-1)$	—	0 or 1
	$H(1-2)$	—	0 or 1
	$\vdots$	—	$\vdots$
	$H(23-24)$	—	0 or 1
Day of the week	D(Sunday)	—	0 or 1
	D(Monday)	—	0 or 1
	$\vdots$	—	$\vdots$
	D(Saturday)	—	0 or 1
Season of the year	S(Summer)	—	0 or 1
	S(Autumn)	—	0 or 1
	$\vdots$	—	$\vdots$
	S(Spring)	—	0 or 1

The use of Probabilistic Graphical Model to consider conditional dependence among the random variables, Bayesian Networks is an example of such a model.

The rule based learning algorithms form another category algorithms such as Decision Tree and Decision Table where the model can be explicitly visualized. Function based learning algorithms (e.g Artificial Neural Networks and Support Vector Machine) tries to learn a non-linear function which tries to map the input data into the output class. Finally, there is another category of learner such as K-Nearest Neighbor which is an instance based learner where the input vector (instance) is compared with all the previously seen cases based on a distance metric to find similar cases seen during training. All these classes of learners are described previously. The choice of the classifier is made depending on each of the above categories.

Concerning the prediction part, the contribution of the work is primarily on finding a generic way of modeling the raw data and proposing a generic architecture able to work with multiple algorithms and database. In this chapter, each appliance is predicted separately, in the following chapter, a multi-label learning approach will consider appliance correlation.

As already presented in the previous section, two methods were used for data representation. The first was by the use of threshold binary (mainly for neural network back-propagation algorithm) and the second by the use of real values (for the other classifiers).

The results of this two methods of representation are compared in Table 8.6 on the prediction of the total light consumption of a house. It presents the performance of prediction for 3 algorithms (neural networks, nearest neighbor and decision tree) and using both ways of representation.

TABLE 8.6 – The performance of prediction for the total light consumption of a house. Two ways of representing the data, for three algorithms are proposed.

Data representation	Neural network algorithm	Nearest neighbor algorithm	Decision table algorithm
Real values	81.72 %	80.22 %	83.10 %
Binary values	83.34 %	81.98 %	82.94 %

From the results of the Table 8.6 and for the simplicity of representation, real values are preferred as they present less complexity. Also, the rules associated with a model example (decision tree) are easier to interpret. Only the Artificial neural networks are only modeled using the orthogonal binary representation in our work.

For a simpler modeling, the past consumptions and the future consumptions are encoded using threshold to binary. Other knowledge such as hours, days, seasons are kept as real value, as summarized below :

Past consumption : takes an integer value between 0 and 1 : 0 for no consumption and 1 when the appliance is started in the considered hour.

Hour : takes an integer value between 0 and 23.

Day : takes a value among 0 and 6.

Season : takes an integer value between 1 and 4.

Month : is added to the knowledge and takes an integer value between 1 and 12.

This periodic knowledge is extracted at the oracle stage which primarily considers the periodicity in the data. But this approach can be extended to add other structures in data such as duty cycles. The results of the approach using various classifiers is propose in the Chapter 9.

Summary

The problem of load forecasting is under study for a considerable number of years but it is new at the appliance level. In this chapter a generic appliance load identification technique is discussed and also some common approaches used at the grid level are summarized. The overall architecture is detailed and one of the ways to model knowledge or information is shown. A portion of the overall architecture in implemented, in the future a comprehensive appliance prediction model needs to be built which is capable of modeling user data and process other contextual data. The idea of using a knowledge generator can be extended further to incorporate different information and model them together. In the following chapter the results using this model is observed for different categories of learners.

Chapitre 9

Residential Load Prediction Results

9.1 The oracle features performance **105**
9.1.1 Classifier comparison on predicting the next 24 hours 106
9.2 Using different input categories and classifiers **107**
9.2.1 Next hour prediction for different appliances 109
9.2.1.1 Prediction accuracy for the energy consumption of the lighting 110
9.2.1.2 Prediction accuracy for the energy consumption of the washing machine . 112
9.2.1.3 Prediction accuracy for the energy consumption of the electric oven . . . 114
9.2.2 Discussion 115

Introduction

In this chapter the results for the generic prediction model are observed, with focus of the impact of various categories of knowledge for the appliances present in the house. The results are organized as follows :

1. First the results obtained after the addition of expert knowledge.
2. Next, the hourly prediction results using various classifiers for different appliances.
3. Lastly, the future 24 hours results are compared with some trivial knowledge driven predictors.

The scoring is done in terms of accuracy, where accuracy is the number of correct classification to the total number of cases.

9.1 The oracle features performance

In this section we observe each of the proposed knowledge fed to the oracle and its impact in terms of performance of prediction. For each appliance in a house we need subsequently to select a subset of the knowledge proposed by the oracle. Therefore, a feature selection method has to be used (for example the *Principal component analysis*).

We randomly choose a house, and then an appliance in that house. The classification performance using different features in sequence is observed. The Neural Network (NN) classifier (section 8.1) is used to predict the following 1 hours in the future. It can be seen from table 9.1 that due to our incremental approach the knowledge which appears first has more impact than the subsequent feature.

In practice, a feature selection algorithm (principal component analysis) is used. The feature are ranked and they are added in the order of ranking. The results indicate that the same appliance (lamp) has different performance based on its usage patterns in two houses. It can further be observed that adding features does not always ensure better performance as the dimension of the feature space is increasing and sometimes it leads to over fitting the training data set. The accuracy measure is high for water heater and washing machine prediction as observed from table 9.1 but the number of OFF cases are also high in these two appliances. Appliance wise results will be further explored in table 9.2.

Table 9.1 – Oracle Result of various appliances based on different features

Knowledge	Appliance				
	Lamp-1	Water Heater	Telev. set	Lamp-2	Washing Machine
Past consumption	82.94	94.87	89.10	78.90	84.95
Time of the day	83.45	95.00	89.84	81.55	85.42
Day of the week	83.73	95.00	90.03	81.41	86.02
Season in the year	84.14	95.10	89.84	81.276	85.79
Same hour of the previous 7 day	83.50	95.05	90.17	81.27	85.74

9.1.1 Classifier comparison on predicting the next 24 hours

After the selection of the data, we predict the following 24 hours at a sampling rate of one hour. The results in terms of prediction accuracy are proposed in the Table 9.2. The model for the future 24 hour prediction is developed by treating each of the 24 hours separately. The features are adjusted accordingly with the past consumption only available from the hour H (current hour) and before. In the implementation the prediction of hour $H{+}1$ is not used for the prediction of hour $H{+}2$ and so on. This is to avoid error propagation.

The prediction system is evaluated with two methods. The first is by simple averaging all the accuracy for the 24 hours. The second is by using a weighted average. The proposed weighting scheme is expressed by the following equation :

$$\sum_i \frac{2.(24-i)}{25*24} * Accuracy(i) \text{ for } i \in \{0, 1, \cdots, 23\} \tag{9.1}$$

The weighting scheme gives the highest priority to the nearest time interval and the least to the last one. The predictor is learned every hour, so the last hour ($H{+}24$) will be predicted 24 times before being observed. The table 9.2 and the figure 9.1 compares the Neural Network classifiers (NN) with other trivial classifiers. The trivial knowledge based predictors are defined as following :

- The predictor that always predicts the appliance won't start is called *never starts*.
- The predictor that always predicts the appliance will start is called *always start*.
- The predictor that always predicts "what happens the previous day at the same hour happens the next day" is called *24 hour similarity*.
- The predictor that always predicts "what happens the previous week at the same hour happens the next day" is called *168 hour similarity*.
- The predictor that always predicts "what happen a random hour back happens the next hour" is called *random hour similarity*.

TABLE 9.2 – Comparison between trivial and Neural Network Learner

Appliance	Results						
	Never Start	Always Start	24h simil.	168h simil.	Rand. h. simil.	NN aver.	NN Weighted
Lamps 1	43.73	56.26	71.25	66.99	50.36	78.25	78.85
Electric heater	71.76	28.23	94.38	89.39	92.35	96.31	96.68
Lamps 2	32.02	67.97	88.014	83.35	80.33	90.55	90.82
Electric oven	84.46	15.53	80.32	80.76	72.59	86.00	86.03
Television set	72.13	27.86	71.36	69.43	57.30	76.64	77.18
Electric cooker	92.15	6.05	89.72	89.68	88.80	94.18	94.21
Washing machine	88.88	11.11	86.25	86.79	78.45	89.55	89.89

It the figure 9.1 the complementary trivial predictors *Never start* and *Always start* are grouped together and highest score is presented. Any prediction method should be above this performance level to be useful for the building energy management. These estimates give an overall idea of the performance of our proposed model. The results indicate that the proposed model works better than other trivial predictors. Previous works on appliance usage prediction from consumption data relied heavily on the assumptions expressed in the trivial knowledge based predictor. The assumption of 24 hours or 168 hours similarity is intuitive but other knowledges also need to be incorporated to make the system more dynamic. The trivial predictor provides the benchmark above which the non-linear learner needs to perform. It also needs to be noted that the performance for an appliance is evaluated compared to the benchmark set by the trivial learner.

It is observed that the appliances predictability, is linked to their user behavior. From the results proposed in the Table 9.1, considering the lamps, the applicability of expert knowledge varies not only from appliance to appliance but also from house to house as the user behavior is varying in significant proportion.

9.2 Using different input categories and classifiers

Several experiments were conducted with the intention of evaluating the performance of various categories of classifiers and input features (knowledges) for each of the appliances in a house.

In order to test the accuracy of the prediction methodology, several experiments were conducted in three different situations :

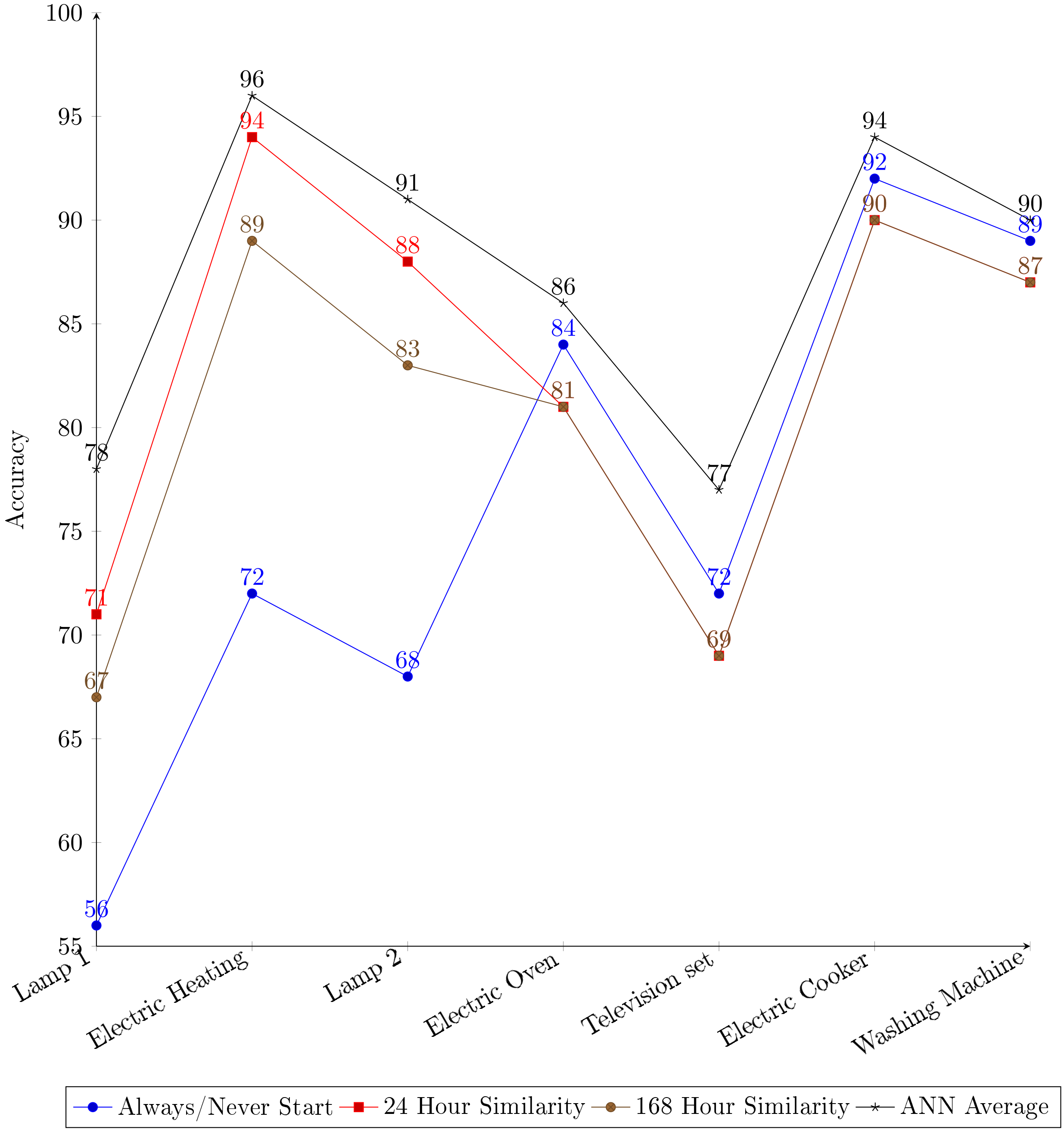


FIGURE 9.1 – Comparison between Neural network classifier and trivial learners

1. Considering *all information*. In that case, we consider the time knowledge from the oracle (hour, day, month and season) and additional information concerning sequential data from the previous occurrences of consumptions.
2. Considering the *time information only*. In that case, just the time knowledge from the oracle is taken into account.
3. Considering the *consumption information only*. In that case, just information from historical data are taken into consideration.
4. Comparison between non-linear classifiers.

The presented classifiers were tested for different periods of historical data : one day (24 hours), two days (48 hours) and one week (168 hours). The accuracy of prediction was computed regarding the state ON/OFF of an electrical appliance for the next hour. The choice of these classifiers is primarily open classifiers whereby the built model can be observed. It is important in this work to exploit the built model to understand the underlying pattern.

From the IRISE database, several houses and appliances are tested with the proposed prediction model. One example of each class of appliances in a house (as defined in the Chapter 7) will be presented in the terms of prediction accuracy. When using the previously defined classifiers, there are other performance coefficients which are used, as seen in the Section 6.5 of the Chapter 6. In this chapter, the accuracy is used as it is generally used for forecasting. Subsequently in the following chapter other performance indicators are explored.

These appliances where chosen as representative for their class : the oven is a *service that is used randomly*, the washing machine is an *appliance which can be delayed* and the lamp is a *service with a regular consumption*. The results will be presented for one of each class of appliance.

9.2.1 Next hour prediction for different appliances

In this section, the results are highlighted after applying the presented classifiers for different services in a house. It is the validation step of the prediction architecture (refer to the Figure 9.2).

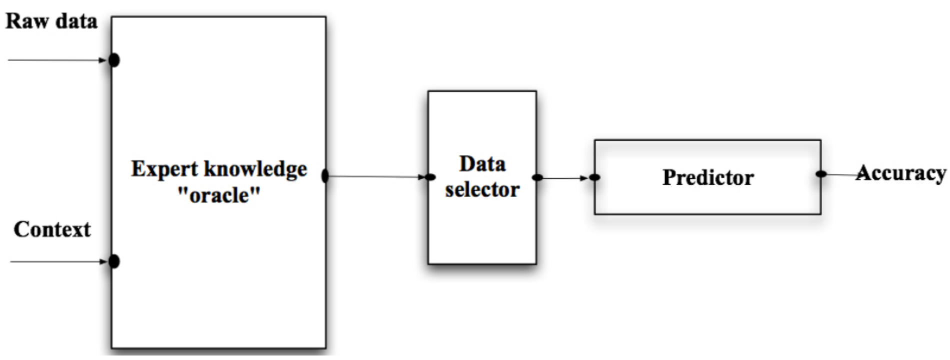


FIGURE 9.2 – A simplified prediction model

since at this point the accuracy of prediction is computed and compared for different classifiers.

9.2.1.1 Prediction accuracy for the energy consumption of the lighting

Lighting has a regular consumption in a house due to seasonal day/night succession. The prediction is done with the proposed architecture using three classification methods (Bayes Net, Decision Table and Decision Tree). These selected classifiers have an advantage that the built model is available and open to analysis which is not case with functional classifiers. The model in figure 9.3 is done using the decision tree classification algorithm and using *all information* available through the oracle knowledges. The model of lighting prediction with decision tree (Dt) and using *only time information* is shown in figure 9.4 and the decision tree model without oracle knowledge (*only consumption information*) is presented in figure 9.5. In the figure 9.3 is can be observed that the node of most information gain is the consumption at the current hour for the lighting. Is is followed by the consumption 24 hours prior. The figure 9.4 is a interesting case to analyse. It can be observed that the lighting is always OFF before 3 PM. It has a very low probability to start between 3PM to 7PM. After the condition of the hour of the day it is observed that the season and month are playing an important role in the decision. In the figure 9.5 is modelling just the consumption information. It can be observed that the previous hour consumption and consumption in the same hour previous day is playing a important role in this case.

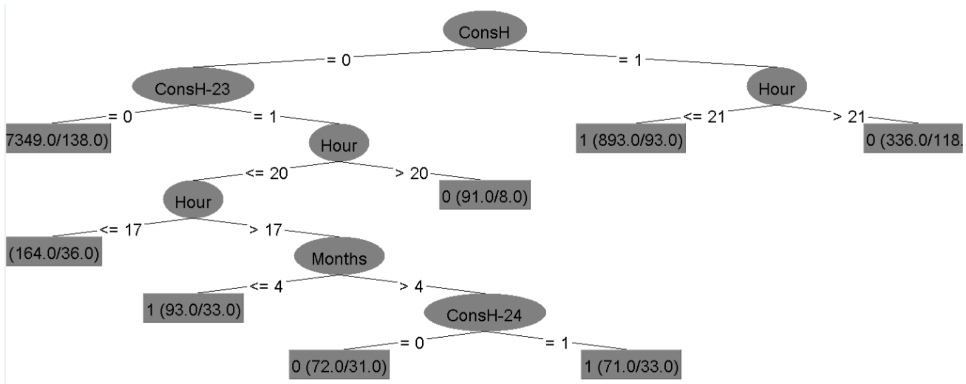


FIGURE 9.3 – The lighting prediction model (using DTL) in the house N° 925 of the IRISE database with *all information* configuration for the oracle.

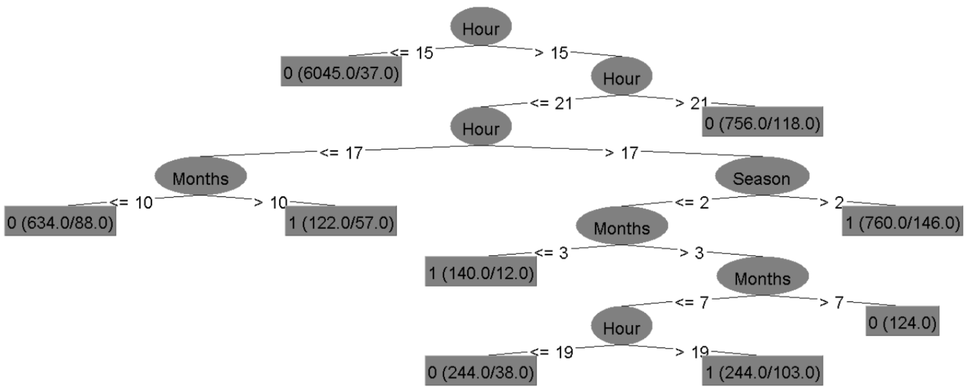


FIGURE 9.4 – The lighting prediction model (using DTL) in the house N° 925 of the IRISE database with *only time information* configuration for the oracle.

In the Figures 9.6 and 9.7, we propose a sensitivity analysis on the prediction results for two classifiers (Decision Table and Decision tree (C4.5 algorithm)). This sensitivity analysis is

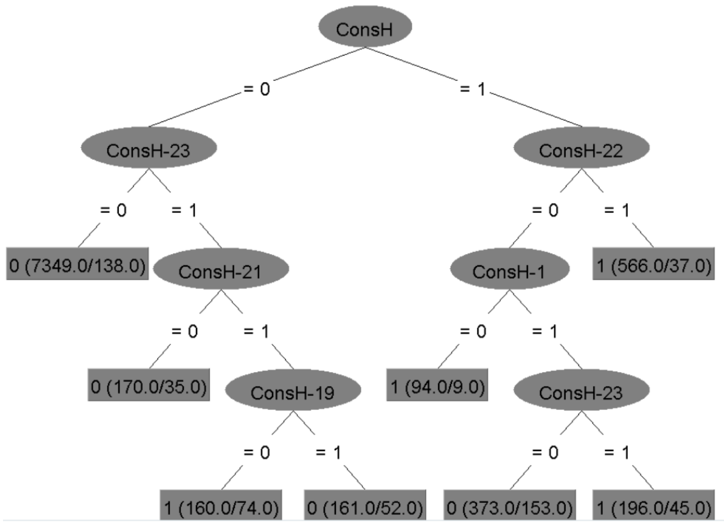


FIGURE 9.5 – The lighting prediction model (using DTL) in the house N^o 925 of the IRISE database with *only consumption information* configuration for the oracle.

conducted on the length of the historical data : from 24 hours (1 day) to 168 hours (7 days).

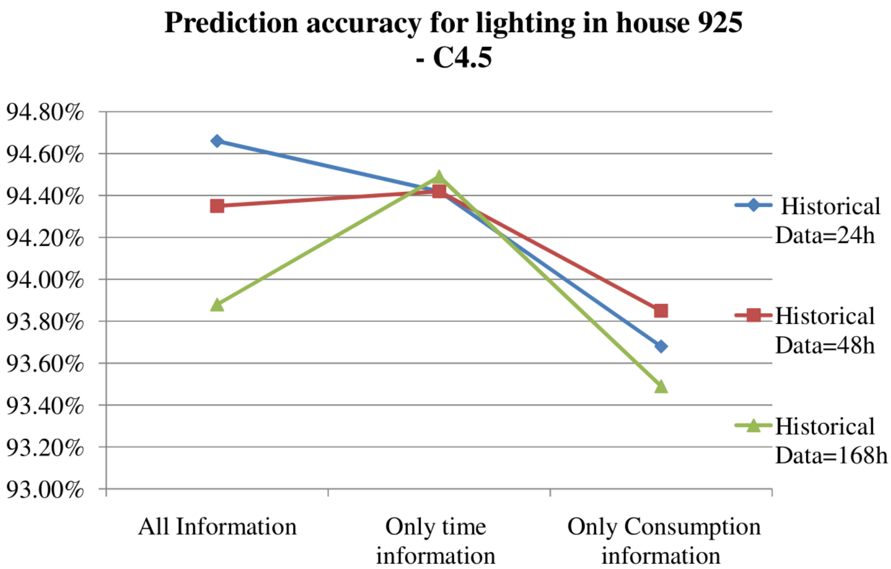


FIGURE 9.6 – The lighting prediction accuracy in the house N^o 925 of the IRISE database using Decision Tree (C4.5 algorithm)

As shown in the Figures 9.6 and 9.7, the maximum accuracy is obtained when using all the available information and the Decision Table classifier. Also, the best results are obtained for a historical period of 24 hours, no matter which method is used.

The result illustrates the fact that depending on the appliance, a massive historical database is not always required. Also, the knowledge of too much historical data can be counter-productive as the model is over-fitted, as seen in the Figure 9.6 in the *all information* case for the historical data of 168 hours.

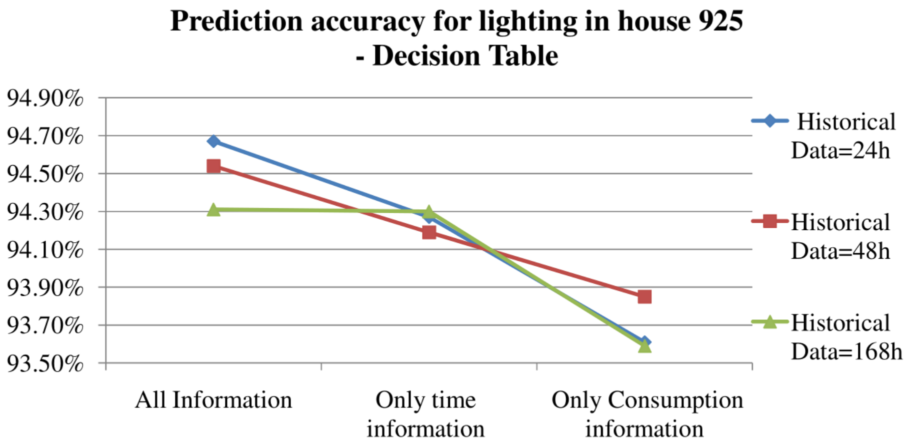


FIGURE 9.7 – The lighting prediction accuracy in the house N^o 925 of the IRISE database using Decision Table (C4.5 algorithm).

9.2.1.2 Prediction accuracy for the energy consumption of the washing machine

The washing machine is an important appliance which energy consumption is interesting to get predicted since it is one of the appliances that can help a home automation system. Indeed, its energy consumption can be delayed.

In the Figures 9.8 and 9.9, we propose a sensitivity analysis on the prediction results for two classifiers (Decision Table and Decision Tree). This sensitivity analysis is conducted on the length of the historical data : from 24 hours (1 day) to 168 hours (7 days).

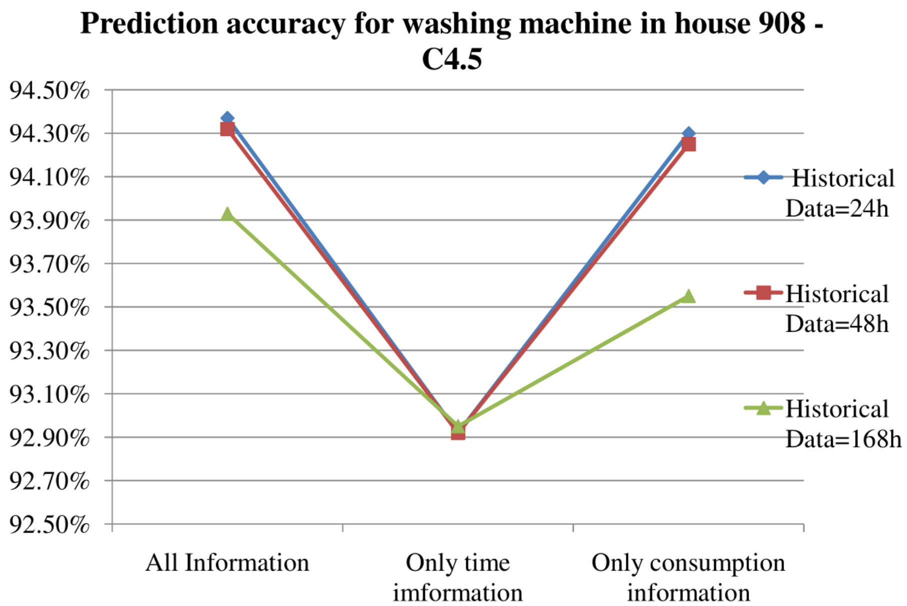


FIGURE 9.8 – The washing machine prediction accuracy in the house N^o 908 of the IRISE database using Decision Tree (C 4.5 algorithm).

In the considered house, the accuracy is the highest for the decision tree method and all the available information from the oracle (both temporal knowledge and information regarding previous consumptions). It can be observed in this case, the results are not sensitive on the length of the historical data. Indeed, this appliance is not known for being rather regularly used in typical households. Therefore, the dependence of the prediction on historical data is weak.

For that reason, the best results are obtained considering all the information or only the consumption information, whatever the classification algorithm being used. The time information is less relevant in the context of the washing machine.

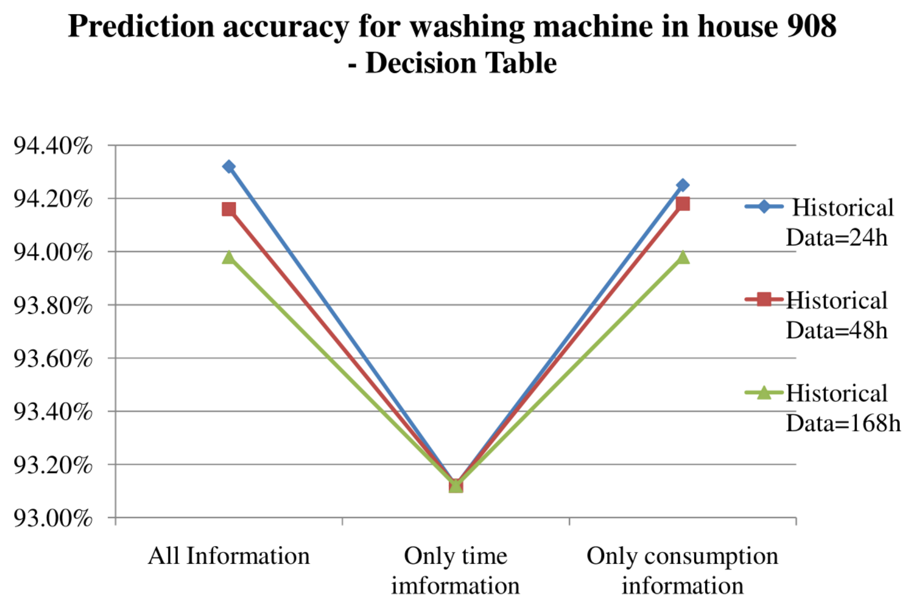


FIGURE 9.9 – The washing machine prediction accuracy in the house N^o 908 of the IRISE database using Decision Table.

We can also observe that the results are very similar for both algorithms.

For another illustration understanding of the results, we propose in the Figure 9.10 the actual predictions for a 24 hours interval compared to the real data for the washing machine in the house N^o 908 of the IRISE database.

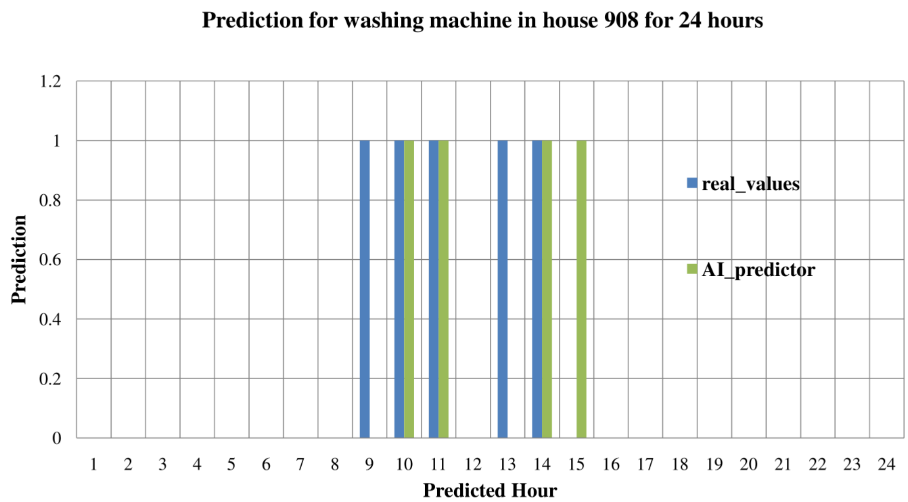


FIGURE 9.10 – The washing machine prediction and real data for 24 hours in the house N^o 908 of the IRISE database.

In this figure, it can be observed that the precision of the washing machine is 60 percent for this particular day. It can be understood from this figure that the error is primarily in the prediction of the first hour of the usage or the start time and the duty cycle for the washing machine. This is an important observation and these features could be incorporated in the oracle to enhance the performance.

9.2.1.3 Prediction accuracy for the energy consumption of the electric oven

In a house, there are several appliances that have random energy consumption. These appliances, very dependent on the user behavior are the electric oven, the microwave, the television etc. The unpredictability of these appliances is not really a problem from the grid point of view, because they are not loads that could represent a potential of flexibility. Indeed, no one will let someone else control there television set, or microwave oven !

From the point of view or the inhabitants, the prediction of these appliances could provide information for a potential change in the way they consume energy, like starting some appliances at another time of the day, for example where the electricity is cheaper.

In the Figures 9.11 and 9.12, we propose a sensitivity analysis on the prediction results for two classifiers (Decision Table and Decision Tree (C4.5 algorithm)). This sensitivity analysis is conducted on the length of the historical data : from 24 hours (1 day) to 168 hours (7 days).

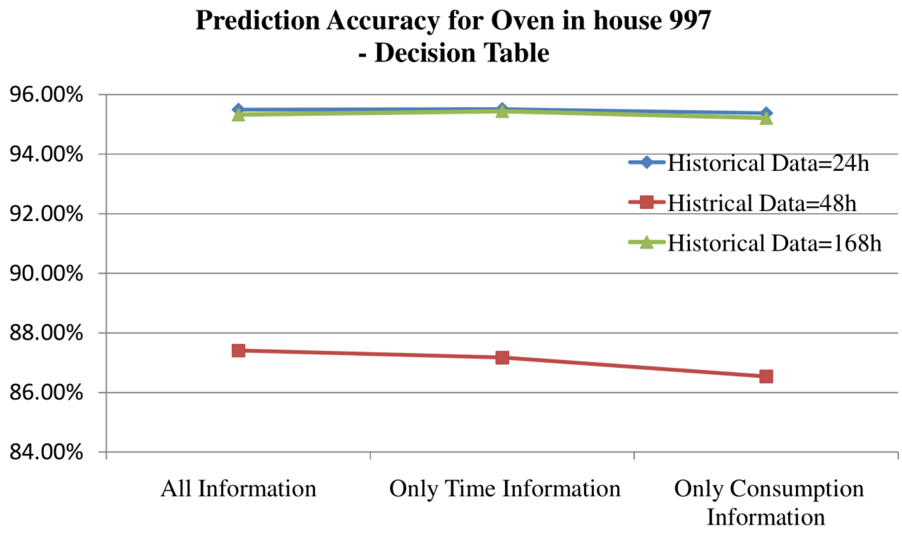


FIGURE 9.11 – The electric oven prediction accuracy in the house N^o 997 of the IRISE database using Decision Table.

Comparing the Figures 9.11 and 9.12, the best accuracy is obtained when the whole information is used, with the Decision Table method and for a historical interval of 24 hours.

For both algorithms, but in a clearer way in the case of the decision tree algorithm, the best results are obtained by providing only time information, and not using consumption information. This observation can be related to the particular behavior of the appliance. Indeed, its usage is not dependent on previous usages, but rather on the time of the day, with a certain variation depending on the way inhabitants every day live their life at home.

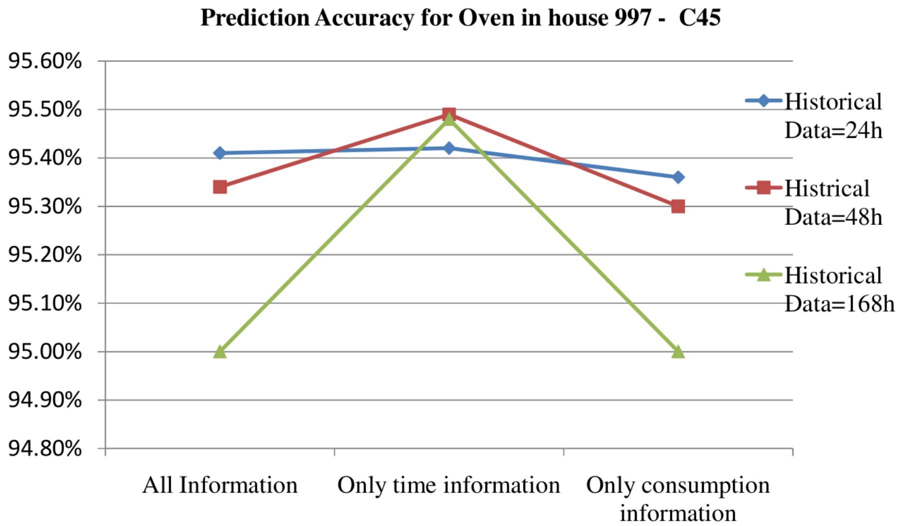


FIGURE 9.12 – The electric oven prediction accuracy in the house N^o 997 of the IRISE database using Decision Tree (C4.5).

9.2.2 Discussion

Many appliances were tested using the proposed prediction architecture. As predictors, four machine learning methods were employed in order to assess their abilities. Although the results for only some relevant appliances were shown, many other appliances in different houses from the IRISE database were tested. It has been noted that in general, the *accuracy of prediction increases with the proposed architecture*, when the prediction considers all the information - both time knowledge and information on last consumptions.

In these experiments, the results are represented by the accuracy of prediction, the index that tells the proportion of the total number of predictions correctly classified. Overall, the global architecture (database, oracle, data selector and predictor) proves to give good results for prediction (the accuracy of prediction is greater than 80 % in all the tested cases). The *oracle knowledge improves the accuracy of prediction.* The accuracy is higher when using all information (which includes the oracle knowledge) than when using just consumption information or just time information.

Results have been divided in three categories, depending on the usage of the appliances :

1. Appliances presenting a regular usage (for example the lights).
2. Appliances which can be deferrable (for example washing machines).
3. Appliances which are used randomly (for example electric oven).

For the first category and third category, the prediction with Decision table taking into account the whole information gives the best results for all the tested houses. But for the second category, the Decision tree algorithm gives the highest accuracy. Although just three examples are shown in the paper, the prediction using this architecture was conducted for almost the entire IRISE database and the results were confirmed.

Conclusion

To anticipate the energy needed for a service in a home automation system, the system must take into account the actions which will be done by the inhabitants. In this context, a proper prediction of energy demand in housing sector is very important. This work focuses on the prediction of the appliance usage in housing because it is a very important problem in the home energy management. The objective is to construct a model able to predict the appliance usage in housing which helps the system to organize energy production and consumption and to decide which appliance will be used at each hour (energy planning). The home automation system, based on the prediction for the next day and on price and comfort indexes can advise the inhabitants to change the time of using a certain utility (e.g. delay the clothes dryer operation for a period).

In this work, the prediction purpose was to see if a particular appliance will be used at a given hour. Standard prediction methods are difficult to use when focusing on given utilities. The proposed approach tries to formalize expert knowledge using predicate functions and, also, find a suitable data structuring for the classifier.

A global architecture for prediction has been proposed. It includes an oracle capable of getting knowledge out of raw data and produce new information and a predictor able to say if a service will consume or not for the next hour. The validation of this prediction methodology is done using the Irise database which contains the consumption records of 100 houses for a period of one year. The used classifiers perform better than trivial ones the structuring of the expert knowledge can still be improved and an on-line learning method needs to be proposed. Indeed, it has been seen in the Section 9.1 that the incremental addition of knowledge from the oracle does not increase sufficiently the performance of the system. The results indicate that this prediction architecture is useful since the oracle provides additional information which leads to more accurate prediction. The oracle knowledge can be increased by additional attributes such as weather data, activity of the occupants [Arghira 2012] or other information about inhabitants obtained through a human machine interface.

The proposed prediction model is able to learn from data, so even if new residents move to the house, the system will re-learn the new user's behavior. Of course, the system will require 24 hours before it has enough historical data for the oracle. It must be mentioned that the system performance will improve with more available historical data of the user. A time based study will be an interesting performance indicator of the system for the future.

Future work involves building a general, fully automated and user interactive prediction system for home automation. Data from users, introduced through a human machine interface could be useful for prediction and easy to be integrated with the model, as the architecture allows additional information to be considered.

Chapitre 10

Appliance Usage Prediction using Multi-label Classifier

10.1 Context **118**
10.1.1 The on-line learning model 118
10.1.2 The learned models 119
10.2 Prediction the state of the loads in different conditions **120**
10.2.1 Impact of the categories of input 120
10.2.2 Using various multi-label algorithms 121
10.2.3 Considering different Houses 123
10.3 Load prediction after load identification **123**
10.3.1 Principle and interest 123
10.3.2 Result : prediction after identification 125
10.3.3 Real-time appliance usage prediction 125
10.4 Prediction algorithm and energy management systems **126**

Introduction

In this chapter, a multi-label classification approach is used (as discussed in the Chapter 6) and applied to real-time prediction context. We will call it *online learning model*. The model discussed in the previous chapter (the Chapter 9) is used on a real-time basis, going further on the use of the consumption of other appliances present in the house and the multi-label transformation of the output.

First, the results obtained by dividing the input into two categories is compared to the classifiers as discussed in the previous chapter. Second, the results for predicting the states of appliances using different multi-label algorithms are shown after describing the base classifier used in on-line prediction. Third, the results for different houses are shown using the Rakel algorithm and decision tree (Dtl) as the base learner. As a reminder, the Rakel algorithm is a multi-label algorithm which works in conjugation with label powerset (Lp) transformation, but is computationally less expensive (for a description of the algorithms, please refer to the Chapter 6). Finally, the results obtained using the load prediction based on a previously conducted load identification are discussed.

10.1 Context

10.1.1 The on-line learning model

A stepwise outline of the future usage prediction implementation is enumerated below. These steps can be cycled on a predefined time period, transforming a punctual prediction algorithm in an on-line learning model.

1. All the appliance loads are obtained as input at a 1-hour sampling.
2. Subsequences are generated using temporal sliding window with a window size of 24 units for all the appliances (as discussed in the Section 5.1 of the Chapter 5).
3. The temporal data (hour,date) and meteorological data (temperature,humidity) are used as input for each next predicted hour.
4. The Multi-label classifier with the generated features as inputs and the states of the high energy appliances as outputs classes is trained and tested iteratively.
5. The model is learned iteratively and is tested in an on-line learning procedure (cycling this process to the first item once again).

The Figure 10.1 describes the principle of the proposed model. At every sampled time instance, it is predicted if an appliance will start in the following hour or not.

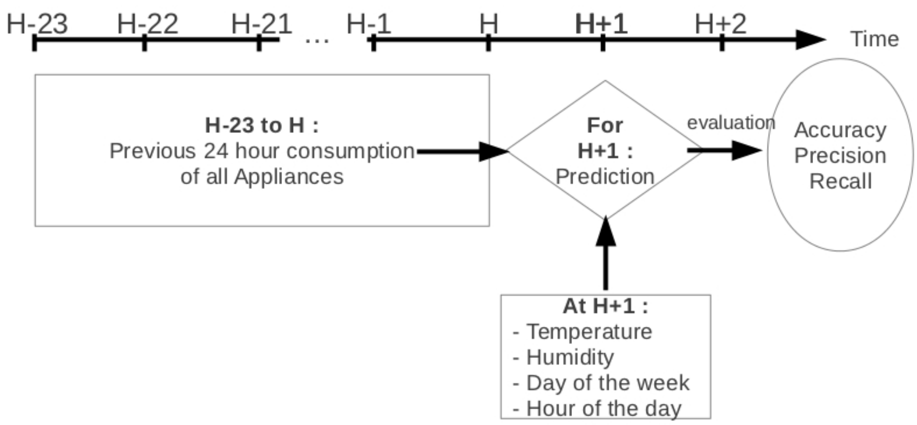


FIGURE 10.1 – Proposed method at a given time instance [Basu 2012a]

The inputs to the model can be categorized thereby :

– The consumption of each appliance at each hour for the previous 24 hours.
– The time of event and meteorological information.

The inputs to the system are shown in the Figure 10.1 for a given time instance (hour H). At each time instance the system predicts the coming hour and then the window is shifted one hour for the next prediction. An important addition has been done in the input from the model discussed in the previous chapter. The input stream consists of the consumption of all the appliances. This gives the classifier the scope to learn inter-appliance correlation within the input stream. In the course of this chapter it is seen that clothes drier consumption depends on the state of washing machine the previous hour. The addition of all appliance consumption results in data dimension overhead and thereby a much higher computation cost.

Two additional remarks have to be made at that point. First, this methodology results in a high amount of data which is processed subsequently at the prediction stage. Second, a significant constraint is that the states of the appliances of the previous hours are sequential. This information is known only after the availability of the previous events. Then, it is only possible to predict the Hour $(H + 1)$ if we have all the needed information about the Hour H. On the other hand the time of event and meteorological information (temperature and humidity) are available for future time instances and are relatively independent of the current time. Furthermore, the time of event is expressed as two periodic variable : hour of the day and day of the week.

This way to consider the time allows to take into account the periodic nature of human behavior. Note that we could have also added the week of the year, or the season of the year, etc. Focusing on a daily prediction, taking too much time related inputs would not have changed the results.

As an illustration, the results for the confidence of the predictor is shown in the Figure 10.2 compared to the actual consumption. The window presenting the different instances will be moving regularly when integrated in an on-line learning predictor.

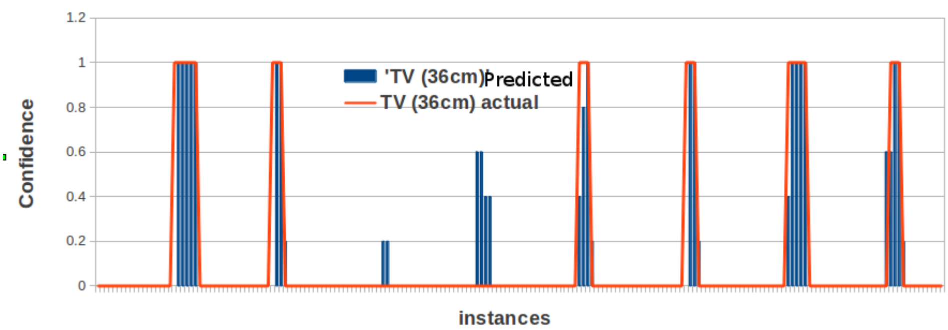


FIGURE 10.2 – Results of confidence of prediction of a television set compared to its actual consumption.

In the classification algorithm, the threshold for the confidence value is determined by the base algorithm. In our case, the confidence value is threshold to "consumption" or "no consumption" (i.e. the state of the load). But in the practical scenario this threshold may not be necessary and the confidence values can directly be used by the energy management system.

10.1.2 The learned models

The learned models gives us an insight into which are the determining factors (inputs) from which the classification algorithms can better predict the future state of the appliances in an house. An illustration of the kind of build models obtained through the chosen classification algorithms is proposed in the Figure 10.3.

This figure is only one example. In fact, the determining factors not only vary between appliances but also vary for the same appliances in different houses. This indicates that the user

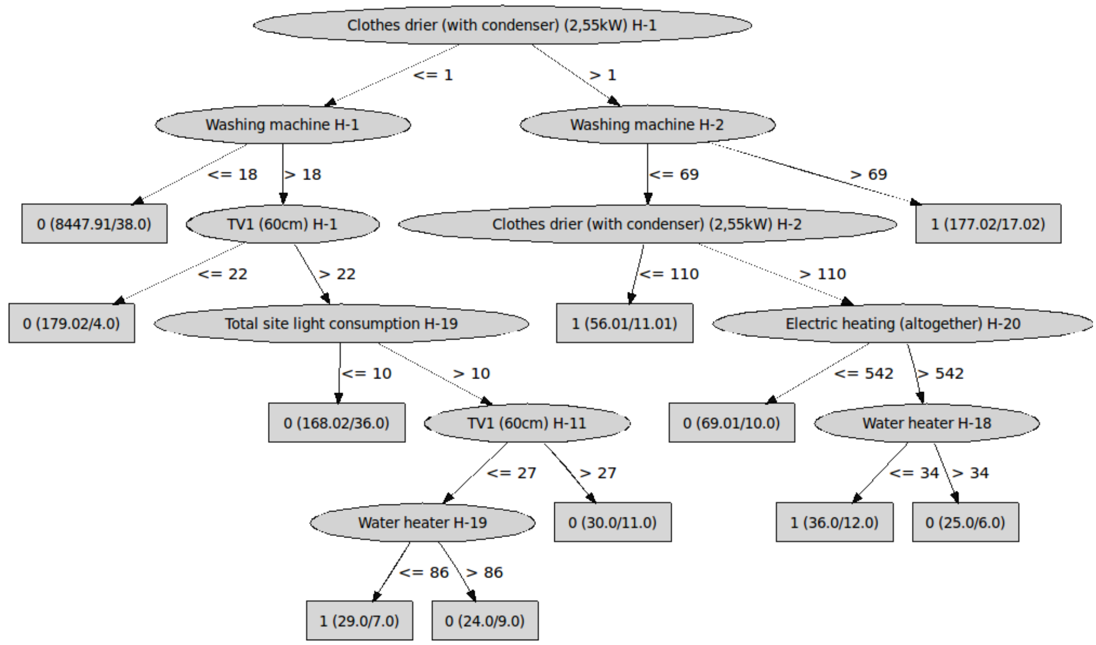


FIGURE 10.3 – Learned Decision tree model of the appliance clothes drier (with condenser)

behavior (i.e. the energy consumption of the appliance) is highly variable and the predictability of the associated appliances depends on the regularity of usage patterns of the inhabitants. The Figure 10.3 shows also that the rules learned for the electric cooker depend on the usage of other appliances services requests. From these built models, a lot can be learned about the behaviors pattern inside buildings.

The Decision tree as presented in the Figure 10.3 consists of nodes where a logical decision has to be made and connecting branches that are chosen according to the result of this decision. The nodes and branches that are followed constitute a sequential path through a decision tree that reaches a final decision. This final decision being in our case the state of the considered appliance. The leaf nodes in the decision tree give the number of instances correctly classified by the built model. For example $0(x, y)$ can be explained as : x is number of instances correctly classified and y is the number of instances incorrectly classified. Still in the Figure 10.3, it can be observed that clothes drier highly depends on the usage of Washing Machine at the previous hour.

The clothes driers are not the easiest appliances to predict. Some appliances such as a television set and the lights are more correlated to hour of the day and even temperature. Therefore, such appliances can easily be predicted for much more hours in the future.

10.2 Prediction the state of the loads in different conditions

10.2.1 Impact of the categories of input

The two categories of inputs that have been discussed in the previous section (one depending on the time, the other on the energy consumption) and illustrated in the results of the Chapter 9 can be also used separately in the predictor algorithm in order to assess their impact on the results. In the Table 10.1, the results obtained for these two categories of input are shown.

TABLE 10.1 – Accuracy of the prediction with two different kind of inputs

Appliance	Qualification	Inputs	
		Only Temporal Input	Sequential and temporal input
Microwave oven	Precision	68.24	68.30
	Recall	15.38	23.30
	Accuracy	85.74	86.40
Electric Cooker (hot plate+oven)	Precision	62.16	68.06
	Recall	5.73	20.19
	Accuracy	95.36	95.76
All lights	Precision	69.47	78.77
	Recall	76.54	79.18
	Accuracy	73.19	80.21
TV (36cm)	Precision	56.37	83.96
	Recall	61.52	72.86
	Accuracy	91.29	95.84
TV (56cm)	Precision	51.32	71.99
	Recall	47.73	68.41
	Accuracy	76.07	85.72
Washing Machine	Precision	95.74	92.02
	Recall	0.0	48.05
	Accuracy	0.0	97.61

These results give a general idea of the appliances that can be predicted by temporal information only and the appliances that are more correlated to the consumption sequence of the past 24 hours. For example, the usage of the lights and of the television set in the considered house of the IRISE database are more dependent on temporal information and therefore may be predicted just based on these kind of information as input (such as hour of the day and day of the week). It must be mentioned that this remark is not true for all houses of the database and varies also from one appliance to the other.

A generic methodology of appliances usage prediction for private houses should then have the ability to adapt the computations to the variability of the loads profiles and inhabitants behaviors. A learning algorithm is, from our point of view, the proper class of algorithm to consider.

10.2.2 Using various multi-label algorithms

In the Table 10.2, the presented results were obtained using the multi-level learner algorithms (discussed in the Chapter 6) on all the appliances. For a reminder on the definition of these algorithms and the explanation of the acronyms, please refer to the end of the Section 7.3 in the Chapter 7.

TABLE 10.2 – Prediction results using different multi-label algorithms.

Appliance	Measure	Algorithms acronym				
		BR1	LP	CC1	CC2	MLK
Water Heater	Precision	57.70	60.48	47.48	57.07	60.89
	Recall	44.01	38.95	47.64	49.42	13.47
	Accuracy	74.19	74.71	69.28	74.35	72.17
Electric cooking (Hotplates + Oven)	Precision	98.58	98.41	98.59	98.47	98.22
	Recall	99.12	99.42	99.20	99.29	97.01
	Accuracy	97.76	97.88	97.85	97.81	95.37
Clothes drier (with condenser)	Precision	71.37	74.55	62.03	72.54	75
	Recall	58.30	59.43	61.69	62.53	4.22
	Accuracy	97.41	97.58	96.97	97.57	96.13
Total site light consumption	Precision	78.18	98.41	71.66	75.50	69.79
	Recall	75.15	99.42	70.16	80.46	77.69
	Accuracy	73.61	75.96	68.96	75.74	70.09
Electric heating (altogether)	Precision	96.54	96.72	94.43	96.94	88.16
	Recall	97.64	97.48	95.51	97.72	89.69
	Accuracy	97.58	97.59	95.97	97.78	90.75
TV1 (60cm)	Precision	98.09	98.42	97.81	98.17	94.54
	Recall	96.88	97.44	97.02	97.48	79.98
	Accuracy	95.90	96.63	95.78	96.45	79.82
Dish washer	Precision	90.90	89.31	85.23	95.38	78.37
	Recall	51.50	54.21	54.50	53.21	12.44
	Accuracy	98.60	98.54	98.56	98.71	97.62
Washing machine	Precision	88.30	90.96	70.28	86.50	42.85
	Recall	55.90	56.79	57.14	57.67	7.93
	Accuracy	96.73	96.90	95.74	96.74	93.48
TV 2 (60 cm)	Precision	85.69	81.80	68.59	81	50.52
	Recall	61.61	64.97	57.57	61.61	11.09
	Accuracy	95.28	95.21	93.34	94.88	90.34
Average of all appliances	Precision	86.12	90.96	79.43	86.04	75.35
	Recall	73.96	56.79	74.03	75.91	48.97
	Accuracy	92.55	92.94	91.08	92.86	87.81

From these results it is observed that the *Rakel* algorithm using the *Label Powerset* transformation present better performances. This may be attributed to the fact that the Rakel algorithm takes appliance correlation into account whereas algorithms such as *Binary Relevance* makes prediction separately for each appliance. We can recognize here similar results as what has shown in the Chapter 5 regarding temporal classification. Algorithms proposing appliances correlations present a particular interest when applied to houses where they are some high energy consuming appliances and not much multiple similar ones.

10.2.3 Considering different Houses

Categories of houses have already been defined in the Chapter 4. Taking into consideration the three main clusters, we propose in the Figure 10.4 the average results of the prediction of all appliances in three houses as a comparison (based on previously defined qualification coefficient).

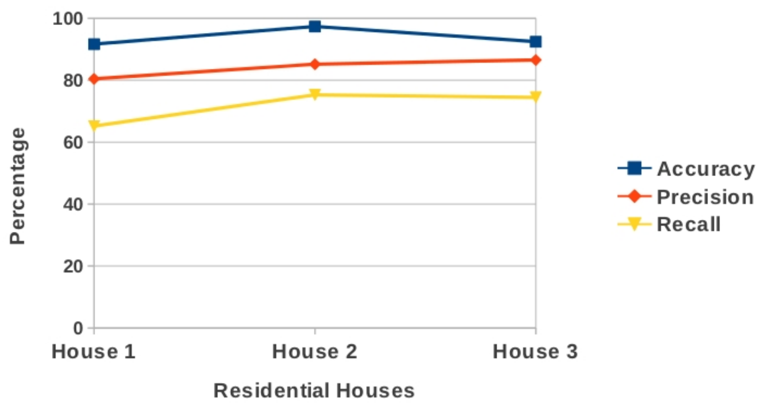


FIGURE 10.4 – Average qualification of the prediction for three categories of houses.

In this figure, the results are obtained by computing the average accuracy of the prediction of all the states of the appliances present in the considered house. This figure illustrates the fact that the same appliance will be more or less easy to process (i.e. classify then predict) depending on the human behavior linked to this appliance. The predictability directly depends on the consumption patterns followed by residents. It can also vary from one day to the other, or one season to the other, etc.

10.3 Load prediction after load identification

10.3.1 Principle and interest

The final objective of this work is to propose an all-inclusive prediction algorithm, capable of predicting the state of the loads without the normally mandatory training phase, including an individual monitoring of the states of the loads (i.e. per appliance). Indeed, the objective would be to be able to predict independently the states or the energy consumption levels of all appliances in a house without accessing these appliances. That means by just monitoring the global energy consumption of the house, gathered at the power meter level.

We already have the identification algorithm presented in part III, the output of this identification being the appliances state only dependent from the consumption of the house (at the power meter). By combining this algorithm and the prediction algorithm, we propose finally a prediction algorithm which does not require the particular knowledge of all appliances in the house during the training phase.

The identification results are used as inputs for the prediction system in an on-line system. The complete work flow is proposed in the Figure 10.5.

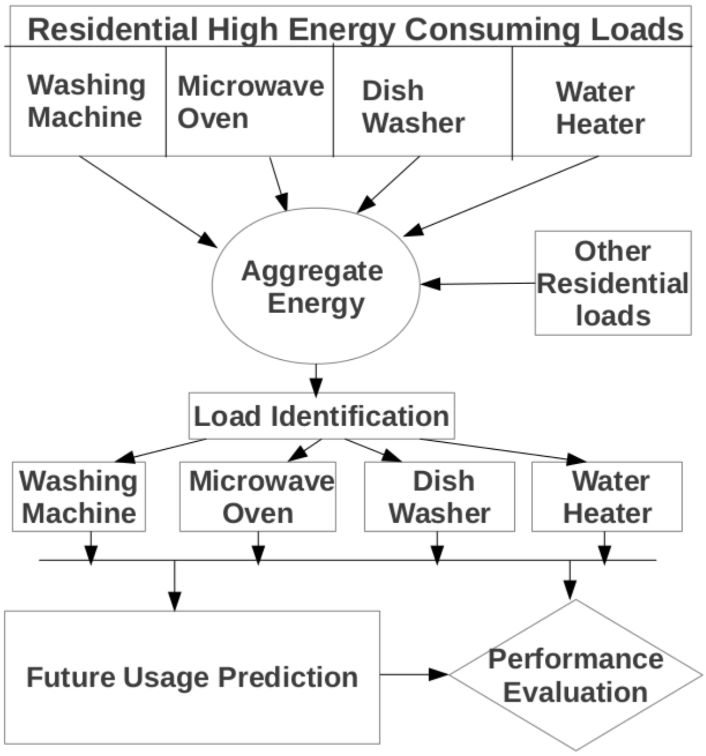


FIGURE 10.5 – Coupled identification and prediction work-flow

A stepwise outline of the future usage prediction implementation based on load identification is conducted as follows :

1. The identified states from the load identification results at 1-hour sampling are obtained.
2. Subsequence is generated using temporal sliding window with a window size of 24 units.
3. Features for the subsequence are generated.
4. The multi-label classifier with the generated features as input and high energy appliances as output classes is trained and tested iteratively.
5. The model is learned iteratively and is tested in an on-line learning procedure.

To summarize, we are actually adding a layer of algorithm of identification to the one of prediction, requiring also a separate analysis of the loads in order for the identification to be effective. But this training phase can be made less intrusive (just conducted once) and requires only time stamped usage indications and not real measurements. Moreover, this phase can be launched periodically in order to re-evaluate the presence of the appliances of the loads in the house in order to detect changes in the number of loads, aging, etc.

10.3.2 Result : prediction after identification

The results obtained by the previously described steps of load-identification-and-prediction is presented in the Table 10.3. These are identification results for a 1-hour sampling rate on one of the 100 houses present in the IRISE database. The identification results are subsequently used as input for the future appliance usage prediction. The sampling rate used here is 1-hour, this differs from the load identification shown previously which was conducted at a sampling rate of 10 minutes.

TABLE 10.3 – Appliances identification and prediction, using the LP algorithm.

Appliance	Qualification		
	Accuracy	Precision	Recall
Washing Machine	95.62	72.80	57.85
Microwave Oven	95.65	14.75	2.88
Water Heater	97.81	89.54	92.80
Dish Washer	97.91	11.90	9.43

The results indicate that water heater and washing machine are identified with higher performance than other high energy consuming appliances in the house. These two loads are interesting for energy management (for example for deferring load) and therefore it is interesting to present good accuracy for this category of appliances. It can be observed, the results are less accurate using the combination of identification and prediction than doing only the prediction with the hypothesis of the availability of training data. It is a logic conclusion (having better results would have been odd) but the difference is still significant. In that configuration, the performance is for some non-controllable appliances, as the microwave oven. Finally the unpredictability of this appliance is emphasized by this specific method. This result could be improved by using a 10 minutes sampling rate.

10.3.3 Real-time appliance usage prediction

Real-time demand response can be complemented if the future usage of deferrable load is predicted with reasonable accuracy after identification. The proposed model takes into account all the possible information based on the identified state of the appliance, the time of the event and the meteorological information. This work uses the learning model presented previously and is launching it again on a regular basis. The input to the learner is the identified state of the appliances presented in the Chapter 7.

The fact that some appliances have high accuracy but low precision and recall (sometimes zero) is due to the dataset being highly sparse and it being representative of only the ON class. Indeed, most of the high energy appliances are OFF most of the time. It shows here a strong relation between the qualification between the two layers of algorithms (identification then prediction).

TABLE 10.4 – Real-time appliance usage prediction based on a previous identification.

Appliance	Algo.	Based on identification (Smart Meter)			No previous identification (direct connection)		
		Accuracy	Precision	Recall	Accuracy	Precision	Recall
Washing Machine	LP	95.11	66.66	18.66	96.58	90.74	64.42
	BR1	95.13	60.51	28.22	96.61	90.00	65.57
Microwave Oven	LP	88.22	13.33	1.41	90.47	32.83	2.75
	BR1	88.27	0	0	90.40	35.92	4.62
Water Heater	LP	95.71	83.42	81.68	98.73	96.29	93.29
	BR1	95.96	86.16	80.33	98.73	96.29	93.29
Dish Washer	LP	95.94	0	0	98.96	83.67	33.60
	BR1	95.94	0	0	99.00	86.00	35.24

10.4 Prediction algorithm and energy management systems

The energy management of residential houses is substantially benefited by the prediction of future usage of appliances. The benefit of using only energy consumption reading for each appliance is many-fold. The objective is to construct a model able to predict the appliances usages in buildings which helps the energy management system to organize local energy production and consumption and to decide which appliance will be used at each hour. The proposed model uses an iterative learning approach that takes into account all the possible information based on consumption data, time of the event and meteorological information.

The learned models are also of great interest regarding the future evolution of this work. The results obtained are encouraging and this model will be incorporated in an energy management system in order to test the viability of the process. Though it must be mentioned that only using energy consumption data as an input has its own restrictions. User inputs to the system can play a major role in improving the overall prediction process.

The results have been validated for a dataset of 100 houses monitored over 1-year. The availability of other datasets could particularly enhance the conclusions of this work. The results suggest that the states of certain high consuming appliances can be identified and predicted even at a low sampling rate of 1-hour. This result is important in the context of energy management and specifically designed for a non-intrusive monitoring of loads and future usage prediction of deferrable loads in the smart buildings and more globally smart grids context.

In the future, load identification at a higher sampling rate may be used to improve the performance of the identification module, with the potential problem of handling critical private information about inhabitants. The future usage prediction is based on identifying patterns in the usage of appliances from the past usage history and then predicting periods when the appliances might be used in the future. The user satisfaction can be taken into account based on these predictions. The practical implementation of the prediction algorithms into an energy management system is discussed in next chapter, the Chapter 11.

Conclusion

The future usage prediction based on iterative learning approach is a logic evolution of the single-use prediction algorithm, taking into account appliances correlation, the identified states, time of event and meteorological information.

The results indicate that the appliances which can be identified with high accuracy and precision can also be better predicted for future usage. It is the case for example of the water heater. As only high energy consuming appliances are considered for prediction, the appliance correlation is not always reflected in the results but still presents interest for the prediction or identification alone.

Both the algorithms give similar performances (for identification and prediction). But the results show a strong indication of dependence between the qualifications of the results. The dataset being highly sparse, only the Off class is properly represented. Indeed, most of the high energy appliances are Off most of the time. Therefore, in order to go further in this work, a new database should be assessed.

Chapitre 11

Case study : Residential Energy Management

11.1 Local Energy Management 130
11.1.1 Definition of the problem 131
11.1.2 System Prototype 131
11.2 The optimal usage of energy storage 132
11.2.1 The objective Functions 133
11.2.2 The constraints 134
11.2.3 The results 135
11.2.3.1 Obj. 3 : Minimizing the bill based on a real-time price 135
11.2.3.2 Obj. 4 : Daily Report peak - effective 137
11.3 Optimal Load Scheduling 138
11.3.1 The objective function 139
11.3.2 The constraints 139
11.3.3 The results 141
11.3.3.1 Load scheduling 141
11.3.3.2 Inhabitants behavior 142

Introduction

In order to validate the use of prediction algorithm, it is interesting to integrate them into an energy management system. As a first step, we propose here different application to what could be conducted on the outputs of algorithms such as the one proposed in the previous chapters.

There are two possibilities. The first one is to consider only the global consumption of a building and try to manage its consumption base on prediction knowledge. The variable here will be an additional component because there is no possible access to any appliance inside the house. In that case, we integrate an energy storage capability as a variable that could be optimized and a local production in order to give more possibilities to the energy management system. The second possibility is to move inside the buildings. Knowing the future state or consumption of the appliances, a local energy management system is able to propose much more than only a storage capability. In fact, without electrical storage, it is possible to act on chosen appliances, eventually thermal storage, in order to propose services to the distribution grid manager (or a future aggregator) or simply to reduce the electricity bill of the inhabitants by acting on their global load curve through deferrable loads for example. Of course, the ultimate smart buildings will include a combination of the two possibilities, with multiple optimization criteria depending on the time of use, prices etc.

In this chapter, the optimization of energy storage and residential loads is first conducted for a prototype building consisting of local production, local storage and residential appliances. The load optimization for various appliances is then proposed using the appliances states prediction algorithms discussed in the previous chapters. The system is tested for various objectives which are beneficial alternatively for the inhabitants and a distribution grid manager, or an potential aggregator.

11.1 Local Energy Management

Based on the methods developed in our research on the prediction of the load consumption, we propose here to test a number of energy usage scenarios between a grid connection, a local production, an energy storage and residential load [Kouveletsou 2012] as summarized in the Figure 11.1. The ultimate goal is to integrate these optimizations results taking into account the uncertainty of prediction, the state of health of the batteries and the different categories of loads. The environmental impact of combining locally these actors are also one of the future directions of the work.

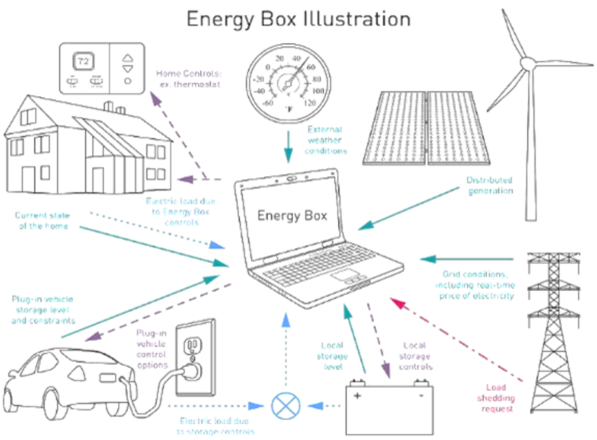


FIGURE 11.1 – Energy Management system for a residential building [Livengood 2009].

Based on a local prediction of the consumption of the total load demand and later the same kind of prediction at the appliances level, we are able to find the best location in time for the use of energy between all these residential actors in order to minimize a certain number of criteria, which are defined in the form of various scenarios. The proposed generic optimization architecture is shown in the Figure 11.2.

Advanced works on the demand-response (taking into account the local environment) requires the consideration of different usage scenarios related to the actors at stakes in the local energy-related supply and utility services [Siano 2014, Berges 2010]. Policies can be simultaneously bound to sources of production and local storage (PV pannels, electric vehicles, etc.), energy efficiency specifications (applications standby, better management of heating, etc.), energy management (charge shift, energy-box, etc.) or specific incentives (variable electricity prices, environmental awareness, etc.) [Palensky 2011, Wacks 1991].

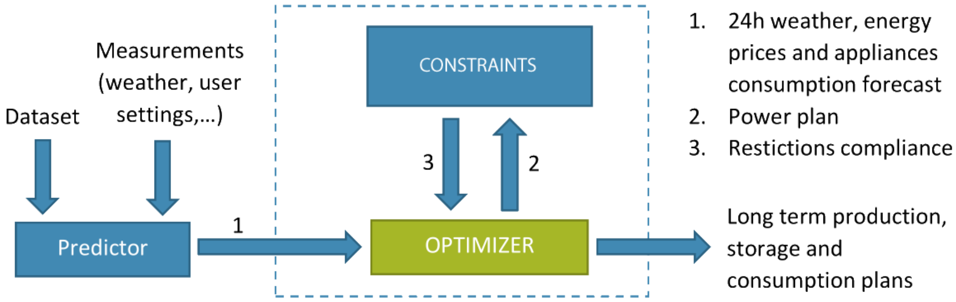


FIGURE 11.2 – Principal synoptic of the considered residential energy management.

11.1.1 Definition of the problem

The results of real time storage and load scheduling optimization of the building are presented for the next day, based on the prediction of the state of the appliances. The local production (PV panel) is assumed to be known.

The parameters of the system are held to satisfy certain constraints defined in the form of usage scenarios. Considering the three phase architecture described in the Chapter 2 , the reactive layer can subsequently correct the value of the system parameters in time to meet a given additional constraints (adapting to a bad prediction, or responding to a specific request from the grid manager).

The considered load curve is the one obtained after the prediction of the global or distinct load curve (passed by both anticipative and reactive layers).

Ultimately, this work will serve as a basis block for analyzing mechanism available to inhabitants and grid managers to act on the new flexibility means offered by a predictive load control. All actions have to be framed by a number of business models developments, that are still to be proposed. Indeed, without financial support (such as incentive price rates for example), there is no reason why people would leave the control of their energy use, agreeing on the one hand to a decrease in their comfort and secondly to an intrusion into their privacy.

Finally, we only consider the energy exchange in terms of electricity. The thermal evolution of buildings, which has an impact on the requirement for heating, may be taken into account by the algorithm in future work. This work was implemented in Matlab in collaboration with two master intern students, William Leon and Ioana-Raluca Gafton.

11.1.2 System Prototype

The considered system prototype is simplified to a PV panel, the load corresponding to one house and its connection to the grid, all controlled by the addition of flexibility of a lead acid battery and controllable loads, as presented in fig. 11.3.

Two of the main components come from experimental readings : the solar panel and the load curve of the house, the last one being the output of the predicted appliances states. Regarding the battery, we used an implemented model derived from previous work [Riffoneau 2009]. For each system, the choice of the PV power and the size of the batteries is a function of the total

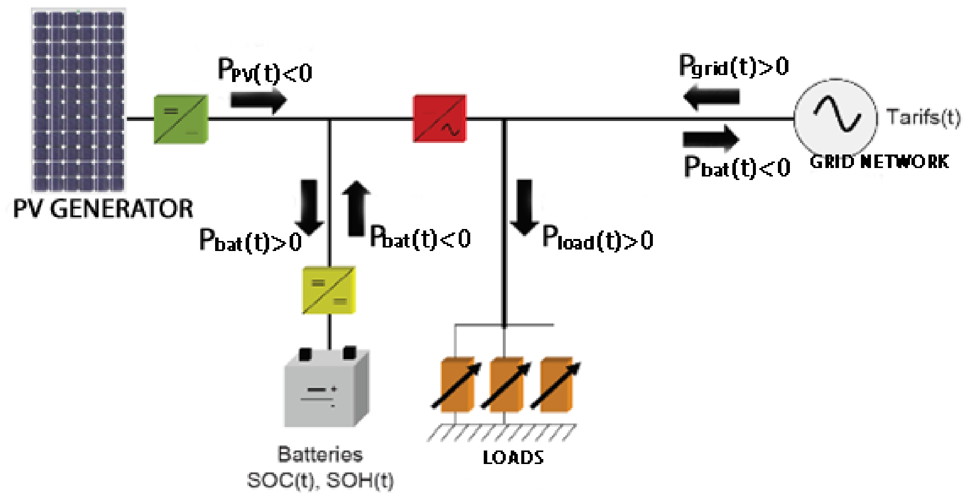


FIGURE 11.3 – Architecture of the considered system [Riffoneau 2009]

consumption of the considered system.

In this work, two main scenarios are considered : the optimal Storage management and the optimal Load Management. The two scenarios are exploited and presented separately in the two following sections, the combined model is for future development.

11.2 The optimal usage of energy storage

Various types of actions can be taken once the physical system shown in the Figure 11.1 is established. Indeed, it is possible to take into account the prediction of consumption, production, the state of charge of the battery of energy costs, etc. [Badreddine 2013, Constanzo 2011]. The characteristics of the components used in this chapter are adapted to the global consumption of the considered house.

The aim of the present work is to provide a block of optimization software used thereafter based on responses to the prediction algorithms. The predictive layer will schedule the controllable loads based on a prediction for the next day, and the other variables in accordance with the later defined objectives. Indeed, it can be a willingness of people to reduce their electricity bill, or an action of the grid managers to ensure the smooth operation of the distribution system of electricity or overcome some technical constraints. These two objectives can be contradictory.

From a more economical point of view, the optimizations based on price criteria (to be used in real time) are using values obtained from the websites of EDF and RTE, as shown in the Figure 11.4. In this model, only the buying price is considered, there is no possibility of selling to the grid. A different price model is considered with both the selling and buying price in the Section 11.3.

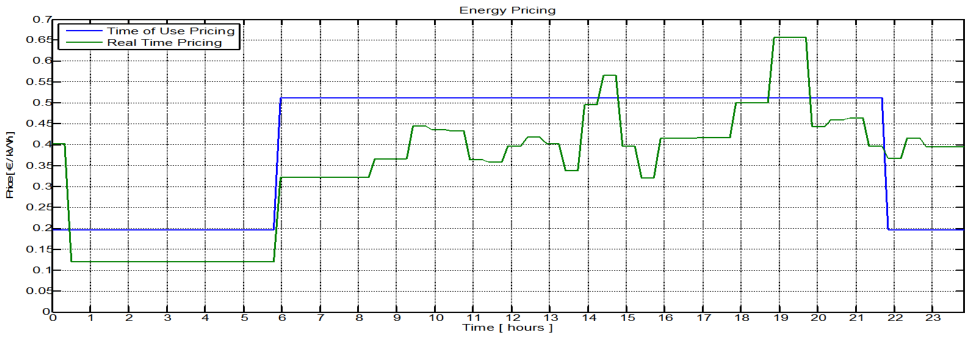


FIGURE 11.4 – Electricity cost : C_{ToU} for time of use (in blue) and C_{RT} for real time (in green).

11.2.1 The objective Functions

The *trivial* or non-optimized case is when the energy is stored in the battery whenever there is an excess of production compared to the load demand and when the battery is discharged whenever there is a load demand and sufficient energy (constrained by the state of charge). This trivial case is used for comparison with all other proposed optimal solution.

We want to optimize the energy exchange between the different components of the considered system. Classically, this optimization problem is based on a cost functions called F which is minimized in the form of the Equation (11.1).

$$\min_x \sum_{t=1}^{T} F(t) \tag{11.1}$$

Different objective functions can be proposed when considering the system shown in the Figure 11.3 [Thiaux 2010]. Two main possibilities can be defined. Either we consider the local grid manager, either we consider the inhabitants. For both, the objectives can be described in term of energy, or in term of cost, both possibilities not being equivalent.

The different objectives, expressed in terms of costs functions, can incorporate economic criteria (energy prices) and technical criteria (energy consumed in the grid, the form of the load curve, etc.). Five examples of objective functions that could be used considering the optimal usage of electric energy storage are describe in the following :

Obj. 1 : Minimize energy exchanges with the grid.

$$\sum_{t=1}^{T} E_{grid}(t) \tag{11.2}$$

Obj. 2 : Minimize the end user cost for a price of use.

$$\sum_{t=1}^{T} E_{grid}(t) * C_{ToU}(t) \tag{11.3}$$

Obj. 3 : Minimize the bill for a real-time price.

$$\sum_{t=1}^{T} E_{grid}(t) * C_{RT}(t) \tag{11.4}$$

Obj. 4 : Minimize the ratio of the maximum and average energy per day, so the fluctuations in consumption are reduced.

$$\frac{\max\left(E_{grid}(t)\right)}{E_{rms}^{grid}} = \frac{\max\left(E_{grid}(t)\right)}{\sqrt{\frac{1}{T}\sum_{t=1}^{T} E_{grid}^2(t)}} \tag{11.5}$$

Obj. 5 : Minimize the difference between maximum and average power per day, so smooth the consumption curve.

$$\left|\max\left(E_{grid}(t)\right) - \frac{1}{T}\sum_{t=1}^{T} E_{grid}(t)\right| \tag{11.6}$$

It can be observed that the various objectives are implemented as single objectives but these objectives could be implemented as a multi-objective constraint problem. At the current state of development, the implementation is ongoing.

Each of these optimization goals is relevant for the user or for the grid manager. These objectives are not necessarily mutually exclusive (a good management of the power system is not in contradiction with a lower energy consumption) but that may be true in some cases (a load management based on prices could have an impact on the comfort of the inhabitants).

The first three objectives are based on the user satisfaction. The goal is to limit their bill based on an evolution of the price that may be more or less variable (based on the prices defined in the Fig. 11.4 and named C_{ToU} for time of use and C_{RT} for real time).

The last two objectives are more interesting for the grid manager. Indeed, smooth consumption by several possible means (approximate peak consumption to the average, or minimize the ratio between the peak and rms) is one of the ways to insure a certain stability of the grid.

11.2.2 The constraints

The optimization is performed under the constraints proposed in the Equations (11.7) to (11.11).

$$E_{grid}(t) = E_{load}(t) + E_{PV}(t) + E_{bat}(t) \tag{11.7}$$
$$|E_{grid}(t)| \leq E_{max}^{grid} \tag{11.8}$$
$$SOC(t) = f(E_{bat}) \tag{11.9}$$
$$SOC_{min} \leq SOC(t) \leq SOC_{max} \tag{11.10}$$
$$SOC(T) \geq SOC_{finish} \tag{11.11}$$

$f(E_{bat})$ is calculated by the battery model and is a non-linear constraint to the problem. It is calculated by using a non-linear battery model. The battery model has been taken from previous works [Riffoneau 2009]. The value of the state of charge at each time t cannot exceed the limits imposed in the optimization constraints, including ensuring continuous optimization from one day to the other, forcing at least an identical state of charge at the start and at the end of the day.

11.2.3 The results

All objectives have been implemented using Matlab as support for the optimization algorithms. We chose to show the optimization results only for two objectives. One oriented towards the inhabitants and the other oriented towards the grid manager. Indeed, the results of the optimization are not the main object of this work, but only use the output of the prediction algorithm. The figures presented below are just used here as a proof of concept.

In the following results, the reduction of costs is calculated in simplified way, by computing the price of the electricity paid by the inhabitants, without taking into account the battery or the PV panels. This hypothesis is used more to provide results of illustration of what can be done based on the prediction algorithms than to demonstrate a real economy.

11.2.3.1 Obj. 3 : Minimizing the bill based on a real-time price

The Obj. 3 is defined to minimize the cost of the energy consumption from the grid, knowing that it depends on a variable price. The Figure 11.5 shows the different energy flows between the elements of the system and the Figure 11.6 proposes the related evolution of the state of the optimized charge of the battery compared with a trivial usage of the battery.

A rough estimation of the economy performed on the electric bill in this case is of 56 %, by decreasing the annual cost of the energy consumption to 61 €. The battery is charged when the price of energy is low, then discharged when the price is high, as shown in the Figure 11.6.

The usage of the battery illustrated in this figure is only possible when the future consumption of energy is already known. Indeed, in order to decide to charge the battery at one moment instead of consuming energy from the grid is the direct result of having at least the future states of the appliances, or their future energy consumption.

Note that this usage of the battery will lead to a faster aging, which has not been taken into account in these results.

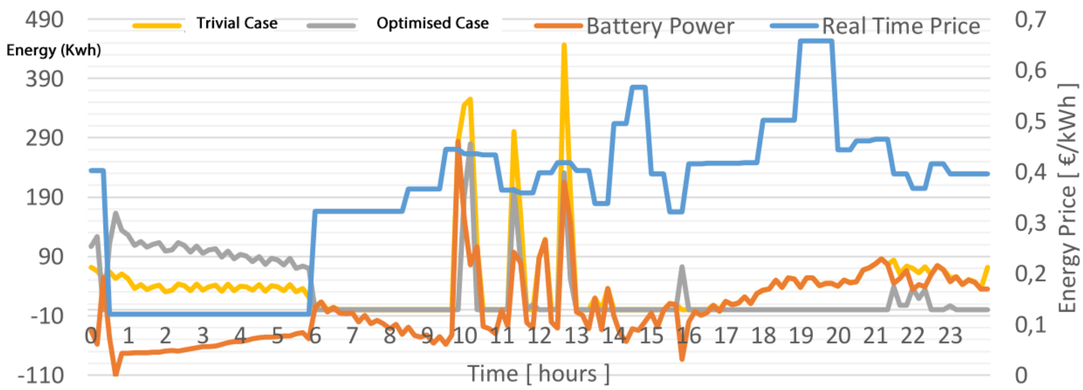


FIGURE 11.5 – Obj. 3 : Optimized energy consumption (grey) compared with the trivial case (yellow).

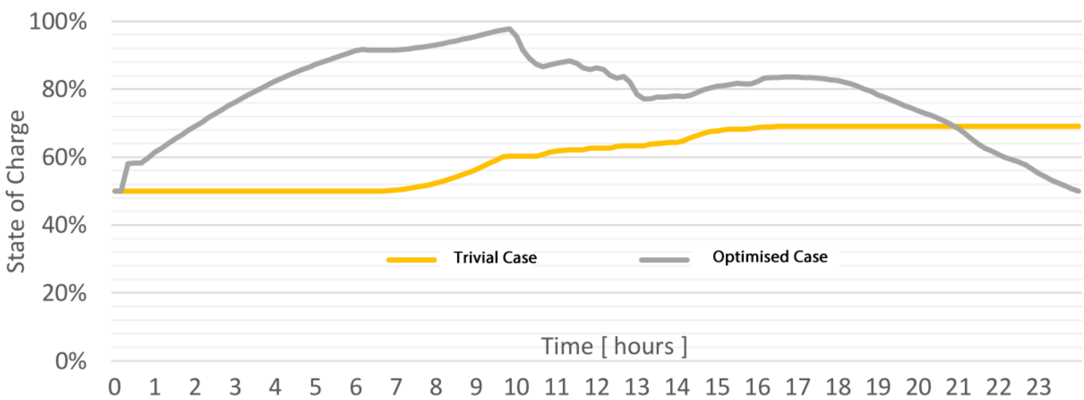


FIGURE 11.6 – Obj. 3 : Optimized state of charge (grey) and the trivial case (yellow).

11.2.3.2 Obj. 4 : Daily Report peak - effective

The forth objective is theoretically designed to help the grid managers to more easily maintain the stability of the grids. In the real life, an average of the consumption of one building is directly made with the adjacent other buildings, leading to more complex problems of optimization. One again, we are only interested here in the concept. The Obj. 4 is then defined as the ratio of the peak consumption and the daily rms consumed energy. This is a way to quantify and minimize the variation of consumption, i.e. to flatten the load curve.

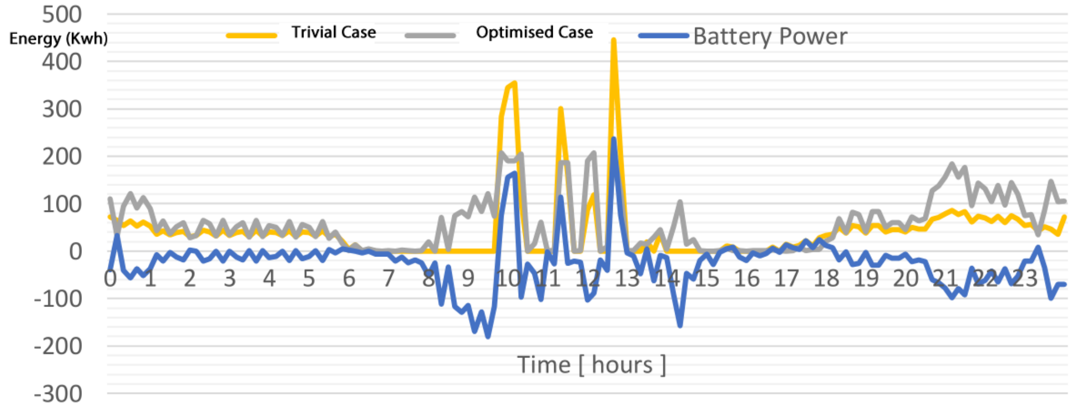


FIGURE 11.7 – Obj. 4 : Optimized Energy Consumption (Grey) and trivial case (yellow).

Without optimization, the coefficient defined in the Equation (11.5) reaches a peak value of 5.63. This value drops to 2.54 after optimization. By cons, such as grid consumption is no longer a goal, we observe an increase of 38 % of it in one day, from 1030 W.h per day to 1422 W.h per day.

The Figure 11.8 shows that the battery is not really requested (it charges almost continuously during the day) because one has to consume more grid power to limit the differences with the peak. Indeed, as no action is taken on the load curve, another component has to make the difference. In fact, the battery is participating in the effort of additional consumption while charging. Not that this effort could not be sustained on a daily basis, because at the end of the day, the state of charge is maximum. A proper use of the battery could be to discharge it during the night for example in a heating system.

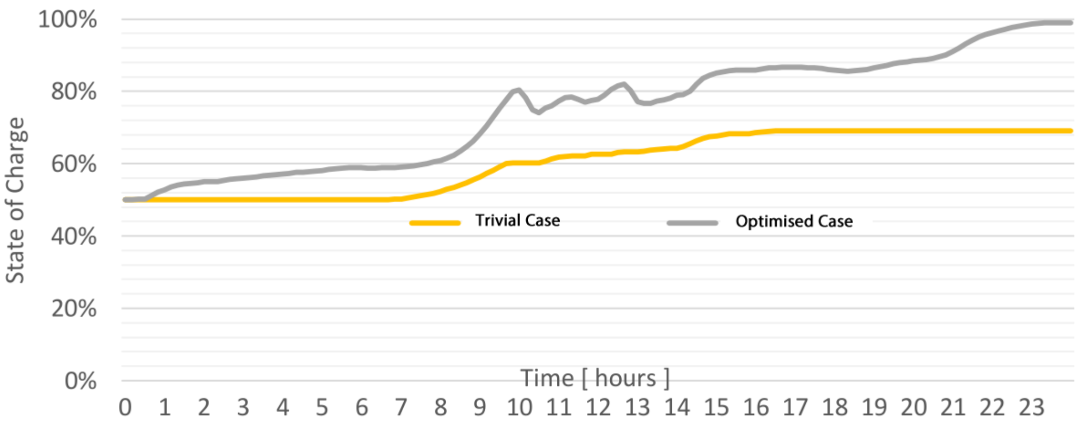


FIGURE 11.8 – Obj. 4 : Optimized state of charge (grey) and the trivial case (yellow).

Analyzing the load curve (for example, using the battery to act on the global load curve of the house) can have a significant impact on the peak coefficient - effective daily. It would subsequently define the categories of charges on which a predictive action is possible and the associated economic model. The use of battery in this case changes the way to interact with the grid.

This particular choice of objective function (limiting the form factor) is made at the cost of not only energy, but also energy bill. To that matter, we have search for a limit case, by removing the constraint of the same level of state of charge at the beginning and end of the day for the battery to improve the coefficient defined in the Equation (11.5).

11.3 Optimal Load Scheduling

The optimized load scheduling highlights the importance of a real-time load prediction system for the grid as well as the inhabitants. Two demand response strategies (linked in their objective) are highlighted as follows :

Price control. Loads control is very sensitive to price fluctuation, especially its augmentation, which at the end will directly influence the residential energy consumption. Speaking about energetic politics, the energy price is used in different manners for different energy sources. Nowadays, the selling price of energy is still relatively fixed and can be divided between full-scheduled and based-scheduled hours. It also has a different balance as a function of the maximum contractual power demanded. In the future, a policy for dynamic energy pricing through an market mechanism capable of handling this scenario is expected.

Direct load control : is a method used by the energy distributors in order to reduce the period of energy consumption, based on the interruptible load control. This technique has been used for a considerable number of years. For example, the water heaters in France are directly controlled by the distribution grid operator through a control signal, based on energy prices. The consumption of certain house appliances (temperature control, water heaters, washing machine, etc.) can also be controlled by the a local energy management which has as main objective to reduce the price of energy production, taking into account the consumer comfort [Missaoui 2014].

To optimize appliance loads scheduling, the considered objective is to minimize the energy cost, assuming that the PV production and the end-user pattern of consumption are known through appliance prediction, but this time without having a storage system. Indeed, the flexibility can now be assumed by the action on the loads. If we can avoid it, no need to pay for another equipment in the house! Without necessitating a non linear optimization solver (mandatory for handling the battery constraints), we have then implemented a linear programming solver for all appliances.

Our implementation in Matlab is based on the *Bintprog* algorithm which solves binary integer linear programming problems, based on the branch-and-bound algorithm (usually suited to solve binary integer problems).

As mentioned earlier, the cost function is directed towards the consumer and its objective is to reduce the overall electricity bill. The prices of energy are then taken into account. In this section, the proposed results are based on the time of use pricing presented in the Figure 11.4.

In France, all the residential photovoltaic production is sold, a selling price being applied in that direction (it is not interesting to consume its own energy). The actual house consumption is covered by the grid connection. Inhabitants are then dealing with two prices, the selling price and the buying price. The bill value C, at the instant of time t is estimated as :

$$C(t) = p_{buy}(t) * E_{tot}(t) - p_{sell}(t) * E_{PV}(t) \tag{11.12}$$

where the total energy consumption E_{tot} is the sum of energy consumptions for each individual equipments, calculated for each time step as :

$$E_{tot}(t) = \sum_{appliances} E_i(t) \tag{11.13}$$

According to [Les Energies Renouvelables 2011], the selling tariffs for April-June 2014 were fixed at 27.94 c€/kW.h. The buying tariff is defined with two prices, fixed from the Blue Contract Option [EDF 2014] at 15.1 c€/kW.h for the hours between 7am and 11pm and 11.77 c€/kW.h for the peak hours (the rest of the day).

11.3.1 The objective function

In the optimal load management, the focus is done on the price optimization. Therefore, we consider the objective function as the sum expressed below :

$$\sum_{t=1}^{m} E_{grid}(t) \times C(t) \tag{11.14}$$

with h the number of hour per day and m the number of appliances in the considered house. This objective function is described in more details below, but determining the different contribution to the global energy consumed by the house :

11.3.2 The constraints

Depending on the category of the appliances, different constraints are elaborated to model their actual behavior.

Interruptible loads. We presume that the total energy consumption of such loads can be rescheduled during one full day into another sum of energy consumption of appliances.

Non-interruptible loads. Their working cycle has to be finished before a requested time imposed by the inhabitants and the sub-phases need to be in series over the considered time horizon.

$$E_{grid}(t) = E_{load}(t) + E_{PV}(t) \tag{11.15}$$

$$E_{Load}(t) = E_{Load_{Controlable}}(t) + E_{Load_{NonControlable}}(t) \tag{11.16}$$

$$\begin{aligned} E_{Load_{Controlable}}(t) = {} & E_{Load_{Interruptible}}(t) * State_{Load_{Interruptible}}(t) \\ & + E_{Load_{NonInterruptible}}(t) * State_{Load_{NonInterruptible}}(t) \end{aligned} \tag{11.17}$$

Also, a non-interruptible load cannot be stopped after having started. This constraints is put in the binary optimization algorithm as follow :

$$\sum_{t=1}^{m} \sum_{i=t}^{D} State_{Load_{NonInterruptible}} = D \tag{11.18}$$

$$State_{LoadIntr}(t), State_{LoadNonintr}(t) \in 0, 1 \tag{11.19}$$

$$E_{grid}(t) \geq 0 \tag{11.20}$$

where,

Indices and set :

t is the discrete time samples having integer value from [1,m].

n total number of appliances.

m total duration of time slot. For one day $m = 24$ is used.

k is the number of non-controllable appliance.

Variables :

$E_{Load_{Uncontrolable}}$ is the energy consumed by the uncontrollable load in the residence.

$State_{Load_{Interruptible}}$ Interruptible load (Water Heater) states (ON/OFF) represented as (0/1) for each time sample. It is a vector of size m, where m is the total number of time slots.

$State_{LoadNoninterruptible}$ Non-interruptible load (Washing Machine, Dish Washer) states (ON/OFF) represented as (0/1) for each time sample. It is a vector of size m, where m is the total number of time slots.

Parameters :

$E_{LoadIntr}$ and $E_{LoadNonintr}$ is the average or maximum energy evaluated from the consumption history of the appliance. It is a vector of size m, where m is the total number of time slots.

E_{grid}, E_{load} and E_{PV} are the Energy demand at the Grid, the total appliance load demand and the energy produced by the Photo Voltaic panel respectively. It is a vector of size m, where m is the total number of time slots.

D is the duration during which a non-interruptible appliance is used. For simplicity, it is the average duty cycle of the appliance(s).

$C(t)$ is the price model.

11.3.3 The results

11.3.3.1 Load scheduling

The results of the simulation are proposed for three controllable appliances. The water heater is considered as an interruptible load, the washing machine and the dish washer are considered as non-interruptible but deferrable loads.

The working cycle of the non-interruptible loads are computed one after the other, as a single block, having as parameters their average energy consumption, allowing to consider simple equivalent "block" of duration of use instead of a real load curve. Without changing the principle of the optimization and the results, this simplification allows to keep a binary optimization algorithm, very fast, in order to illustrate the potential use of the prediction algorithms.

When the appliance is consuming different level of energy, the maximum value is kept. The results are then taken as "worst case scenario" for the optimization.

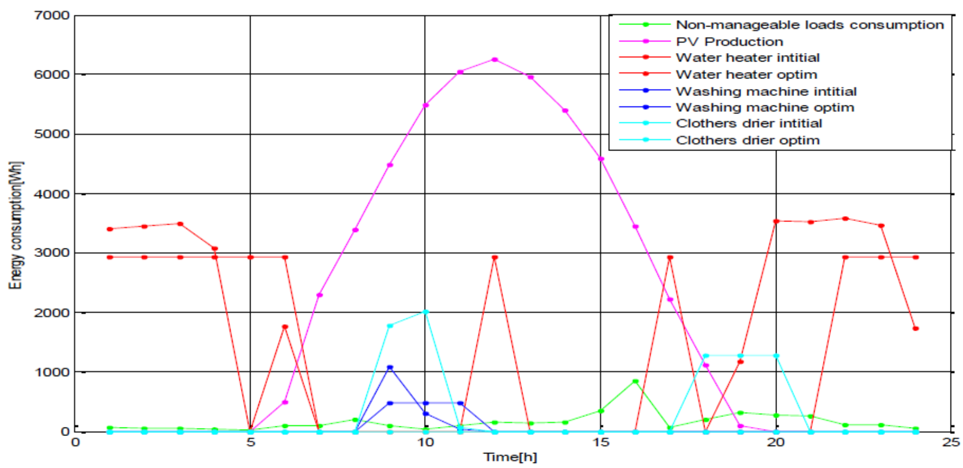


FIGURE 11.9 – Example of load scheduling optimization.

It can be observed in this figure that the washing machine and the clothes drier are used during the OFF hours of the water heater. Without enforcing another restriction, the loads consumption will be shifted as much as possible to the non-peak hours, where the energy price is cheaper.

Also, it is possible to track the difference between interruptible loads (like the water heater in our case) and non-interruptible loads (like the washing machine) depending on the usage of energy. These observations are possible due to the fact that we did not consider any thermal constraints. Therefore, it is possible to move the consumption of water heater and also the heating in the house without considering the evolution of the temperature of the water, or the comfort of any room.

11.3.3.2 Inhabitants behavior

Considering the price mechanism described previously, the Figure 11.10 proposes the reduction in prices depending on three way of scheduling loads in a household.

The three cases are : a trivial one without any load scheduling (i.e. computed directly from the prediction of the states of the loads), a case where there is a limitation imposed to the energy consumption and a case where there is not.

A limit on the daily energy consumption could be representative of some houses that cannot have a permanent grid access. The limitation on the energy could come from a future battery system, installed in order to replace this grid connection in the form of what is called now a micro-grid.

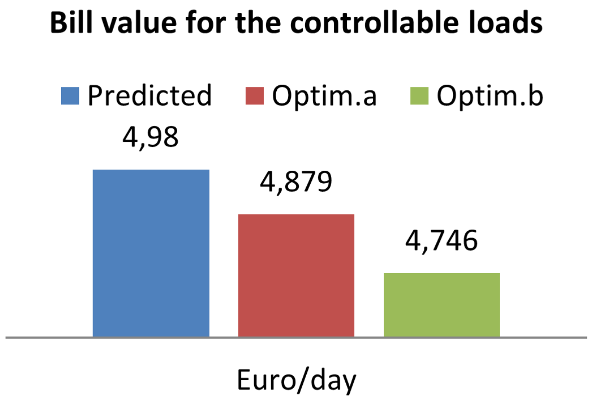


FIGURE 11.10 – Reduction of the electricity bill in an overall of 24 hours based on a prediction, with load scheduling and a limit on the energy consumption (Optim. a) or not (Optim. b).

In the Figure 11.10, it can be observed that there is a decrease in the cost from the trivial to the optimized scheduling case. In the case of no consumption restriction, there is of course the higher possibility of cost savings.

Conclusion

The two perspectives (inhabitant and distribution grid) show wide variations in the use of the battery and a contradiction in the use of energy. This will require a compromise, which will be naturally applied by use of varying prices depending on the needs of the grid manager. Thus, people may be able to accept a degradation of their comfort (e.g., time of use charges imposed) under the condition of reduced energy bills to help manage the grid, or to participate in the effort to integrate renewable energy.

Thereby, without storage association, these goals are difficult to reach. The battery system is not yet considered an option in French homes, except through the use of electric vehicles, which do not at the moment provide all possible actions (no discharge). However, the opportunities offered by the presence of a battery should be taken into consideration in the development of the control of the consumption of housings, towards smart buildings and finally to the smart cities or similar positive energy territories. The optimal load scheduling of the controllable loads within the house based on the predicted appliance usage provides an insight into the constraints of such a system. It can be observed from the results that optimal load scheduling can lead to a lowering of the costs from a consumer perspective but at the price of his comfort.

The use of batteries also adds to the cost not only in an economic manner, but also as environmental impacts. Optimization goals presented here will not have the same environmental impact as they do not use the battery in the same way, so do not have the same action on its aging. To go further in this study, it is necessary to engage in optimization taking into account the full use of the system, that is to say over a much longer period than a day : the complete lifetime of the system, including any maintenance.

Conclusions and Perspectives

Conclusion and Perspective

The work presented in this manuscript proposed various applications for load monitoring and management in the context of residential buildings. The domain of smart grid and smart meter have been studied for some time now but generic and comprehensive mechanism have to evolved for further enhancement. The huge amount of data generated is a challenge and opportunity in this field of research. The various actors present in the energy distribution operation require well established techniques for proper management.

In that context, the thesis first analyses the various challenges to the domain and subsequently searches for generic mechanism regarding potential solutions.

The database of hundred houses IRISE was first analyzed qualitatively to develop techniques for smart meter visualization. A pixel-based visualization is used to observe the data as a whole at various granularity. The database is subsequently clustered in different relevant categories using well established unsupervised learning algorithms. The flexibility of the load demand within the database is also analyzed to evaluate the amount of energy that could be controlled. It is observed that the flexibility in consumption is directly proportional to the usage of high energy consuming deferrable appliances. Although the methodology is simple to comprehend it could have a significant impact in the smart grids control strategies. For the future, energy data visualization tool needs to be developed. This tool should incorporate both smart meter and appliance load analysis. It should also suggest the user remedies for better energy management.

The recent advancement of smart meter data has resulted in a significant amount of consumption data at the grid level. The hurdle is to propose generic methods that could work at the current sampling time and thus preserve user privacy concerns. The work is to design a temporal classification algorithm based on meta-features as inputs. The results are compared with state of the art sequence classifiers such as Hidden Markov Model. A temporal distance based metric is also used in conjunction with the K-nearest neighbor classifier. The above mentioned techniques consider temporal correlation within the input data, which in this case are the smart meter energy readings. The correlation between outputs (appliances) is considered by the use of multi-label learner. It is observed from the results that the algorithms which consider input and output correlation generally outperform the other algorithms. It is also observed that the various appliances have different accuracy of identification. The seemingly easiest appliance to identify is the water heater, and the microwave oven is the most difficult out of the considered appliances.

The proposed techniques of load identification could be used as an application for both the electricity distribution operator and the electricity consumers. The future works in this domain is to build a unsupervised system for load identification working at a low sampling rate and to test it on different scenarios (depending on users behaviors).

In the final part, strategies for load management within a residence based on load prediction are presented. The energy management strategies will be considerably helped by the prediction of individual appliances in addition to the global load curve of the house. This is a challenging task considering that the appliances consumption directly depends on the consumer usage pattern. The approach is first to design a generic expert-based technique for load prediction. It is observed that the usage of appliances in many cases have a temporal and sequential pattern of usage, which is exploited by the system. A number of expert-based classifiers were tested and compared with trivial classifiers. The results indicate that a discriminative classification based approach is performing better than a probabilistic approach.

Subsequently, an on-line learning method is implemented using a multi-label classification approach. This approach is capable of considering appliance correlation through the implementation of a multi-label classification algorithm. Various state of the art multi-label classification techniques are compared and it is observed that the algorithms that consider appliance correlation are performing better than single-class algorithms.

Finally, an energy management system is proposed for both optimal energy storage and load scheduling, depending on the knowledge of the future global load curve of the house, or the future separate states of charge of the appliances. These optimizations require proper market analysis to be beneficial for both the energy distributor and the consumers. In the implementation, various objectives are considered to illustrate the challenges of such a system. The load management indicates that the presence of controllable appliances within the house would result in reduction in peak consumption but the consumer needs to be given incentive to accept such a system. The application is a simple implementation of a more constrained system. Future work in this regard are considering user defined constraints for appliances and confronting the system to the various scenarios of prediction errors. The energy management system could be further improved with the use of on-line algorithms and meta-features.

Table des figures

1.1 Identification and management of appliances through the power meter 8

1.2 The principle of signal separation 8

1.3 Two types of methods for non-intrusive load monitoring 9

1.4 Non-Intrusive Load Monitoring work flow 10

1.5 Non-Intrusive Load Monitoring [Hart 1992] 11

1.6 Classifier architecture : Centralizing information of a smart household. 14

2.1 Actors involved in the Smart Home concept 16

2.2 Home automation principle 17

2.3 Principle of the prediction system 18

2.4 Architecture of a power manager 20

3.1 Comparison of the 100 houses of the IRISE database by average load consumption over one year of data. 26

3.2 Repartition of the consumption of the 100 houses of the IRISE database by average load consumption over one year between all loads and only deferrable loads. . . . 26

3.3 Frequency of usage of different appliances within a house for different energy levels at one hour. 27

3.4 Day to day comparison using pixel based visualization. 28

3.5 Colour Model for Visualization. 29

3.6 1-year representation model of the data with four different granularities. 30

3.7 28 days window size (left) compared to random window size (right). 30

3.8 Monthly energy consumption representation with hourly and daily granularities. . 31

3.9 Yearly Energy consumption with hourly, daily, weekly and monthly granularities. 32

4.1 Number of usages of the appliances for different energy levels. 37

4.2 100 houses projected into two principal axis (using PCA) and divided into 4 clusters 38

4.3 Minimum of flexibility of the group 3. 42

4.4 Maximum of flexibility of the group 3. 42

4.5 Minimum and maximum of flexibility for group 1 (left) and group 2 (right). . . . 43

4.6 Minimum and maximum of flexibility for group 3 (left) and group 4 (right). . . . 43

5.1 Methodology for temporal classification, domains of interest 48

5.2 Data representation in a standard propositional learning format 49

5.3 Graphical definition of some of the main meta-features concerning the aggregated power measurements in a sliding window. 51

5.4 Two dimensional projection of the On/Off states of a water heater using Pca. The pixel size is proportional to the energy consumption. 52

5.5 Hourly usage of the water heater in one of the house of the Irise database over one year. Two states are considered : OFF in blue and ON in red. 53

5.6 Water heater On/Off states projected using Pca for the meta-feature called *distances from the current event to the local maximum and local minimum*. 53

6.1 Construction of a decision tree, applied on the application “electric oven”. 60

6.2 Two time Series having similar values-based characteristics 64

6.3 Two time Series having similar behavior-based characteristics (Temporal Correlation, Tc) 65

6.4 The continuous and discrete variables for the Hidden Markov Model applied to load identification. 67

6.5 Iceberg to showcase the imbalance in the dataset 67

7.1 Hourly time-line diagram, electric cooking (oven + hotplates). The red lines are the identified states of the load and the blue lines are the states from the readings database. 72

7.2 Hourly frequency of usage diagram for an electric cooker (oven + hotplates). . . . 73

7.3 Accuracy as testing time is increased. 73

7.4 Three possible choices for the sliding window position 74

7.5 Cat. 1 : Load identification performances for five different algorithms with a 10 minutes time sampling (algorithms acronyms in Section 7.3). 77

7.6 Cat. 1 : Load identification performances for five different algorithms with 1 hour sampling rate (algorithms acronyms in Section 7.3). 78

7.7 Cat. 2 : Load identification performances considering some grouped appliances with a 10 minutes sampling time (acronyms in Section 7.3). 79

7.8 Cat. 3 : Load identification performances considering some grouped and repeated appliances with a 10 minutes sampling time (algorithms acronyms in Section 7.3). 80

7.9 Load identification subset accuracy for 3 sampling rates and 4 houses. 81

8.1 Short term load forecasting predictor input configuration 91

8.2 Decision table rules used for oven usage prediction 93

8.3 Bayes network for the electric oven in the house N^o 997 of the IRISE database. . . 94

8.4 Correlation between weeks for the total energy consumption of a house over a year. 95

8.5 Correlation between weeks for only the energy consumption of a washing machine in one house over a year. 95

8.6 Statistical probabilistic representation of the consumption of an electric oven (house N^o 997 of the IRISE database). 96

8.7 Prediction architecture 97

8.8 Simple prediction architecture with no expert knowledge addition. 97

8.9 One day ahead prediction for the complete lighting of a house. 98

9.1 Comparison between Neural network classifier and trivial learners 108

9.2 A simplified prediction model 109

9.3 The lighting prediction model (using DTL) in the house N^o 925 of the IRISE database with *all information* configuration for the oracle. 110

9.4 The lighting prediction model (using DTL) in the house N^o 925 of the IRISE database with *only time information* configuration for the oracle. 110

9.5 The lighting prediction model (using DTL) in the house N^o 925 of the IRISE database with *only consumption information* configuration for the oracle. 111

9.6 The lighting prediction accuracy in the house N^o 925 of the IRISE database using Decision Tree (C4.5 algorithm) 111

9.7 The lighting prediction accuracy in the house N^o 925 of the IRISE database using Decision Table (C4.5 algorithm). 112

9.8 The washing machine prediction accuracy in the house N^o 908 of the IRISE database using Decision Tree (C 4.5 algorithm). 112

9.9 The washing machine prediction accuracy in the house N^o 908 of the IRISE database using Decision Table. 113

9.10 The washing machine prediction and real data for 24 hours in the house N^o 908 of the IRISE database. 113

9.11 The electric oven prediction accuracy in the house N^o 997 of the IRISE database using Decision Table. 114

9.12 The electric oven prediction accuracy in the house N^o 997 of the IRISE database using Decision Tree (C4.5). 115

10.1 Proposed method at a given time instance [Basu 2012a] 118

10.2 Results of confidence of prediction of a television set compared to its actual consumption. 119

10.3 Learned Decision tree model of the appliance clothes drier (with condenser) . . . 120

10.4 Average qualification of the prediction for three categories of houses. 123

10.5 Coupled identification and prediction work-flow 124

11.1 Energy Management system for a residential building [Livengood 2009]. 130

11.2 Principal synoptic of the considered residential energy management. 131

11.3 Architecture of the considered system [Riffoneau 2009] 132

11.4 Electricity cost : C_{ToU} for time of use (in blue) and C_{RT} for real time (in green). 133

11.5 Obj. 3 : Optimized energy consumption (grey) compared with the trivial case (yellow). 136

11.6 Obj. 3 : Optimized state of charge (grey) and the trivial case (yellow). 136

11.7 Obj. 4 : Optimized Energy Consumption (Grey) and trivial case (yellow). 137

11.8 Obj. 4 : Optimized state of charge (grey) and the trivial case (yellow). 137

11.9 Example of load scheduling optimization. 141

11.10 Reduction of the electricity bill in an overall of 24 hours based on a prediction, with load scheduling and a limit on the energy consumption (Optim. a) or not (Optim. b). 142

11.11 Architecture de la classification par l'algorithme d'apprentissage 162

11.12 Fenêtre de calcul et caractéristiques 164

11.13 Synoptique d'identification des charges 165

11.14 Principe de fonctionnement de l'algorithme prédictif 166

11.15 Identification d'état des charges (pas de dix minutes) 167

11.16 Identification d'état des charges (pas d'une heure) 168

11.17 Diagramme de fréquence horaire, four électrique (plaques et four) 169

11.18 Confiance de prédiction d'utilisation d'un téléviseur sur une semaine 169

11.19 Principe d'association des algorithmes d'identification et de prédiction 170

A.1 An example of the raw dataset screenshot 179

A.2 An example of the raw dataset screenshot 179

A.3 An example of the raw dataset screenshot after feature extration 180

A.4 Parallel coordinate plot for the IRESE dataset (All the houses) 180

A.5 Parallel coordinate plot for the IRESE dataset (cluster centre) 180

A.6 The water heater usage identification is shown 181

Liste des tableaux

4.1 Features for residential data analysis. 36

6.1 Example of multi-label classification 58

6.2 One of the tables obtained with a Br transformation from the Table 6.1 59

6.3 Example of Lp transformation from the Table 6.1 59

6.4 Confusion matrix 69

7.1 Selection of sliding window orientations. 74

7.2 Impact on the results of using "Energy variation between time steps" or only "Energy" as features for the classification algorithm. 75

7.3 Cat. 1 : Comparison of different multi-label algorithms for identifying the loads of a residential building. 77

7.4 Cat. 2 & 3 : Comparison of different multi-label algorithms 80

7.5 Comparison of Multi-label learners and Hidden Markov Model with Auc measures 82

7.6 Comparison of distance based metric and Hidden Markov Model with Auc measures 82

7.7 Appliance level performance with F-measures (N.A. is used for "Non Applicable"). 83

8.1 Artificial Neural Network parameters chosen in these works. 92

8.2 Orthogonal representation of the hour of the day. 100

8.3 Orthogonal representation of the day of the week. 100

8.4 Orthogonal representation of the season of the year. 101

8.5 The overall oracle output. 102

8.6 The performance of prediction for the total light consumption of a house. Two ways of representing the data, for three algorithms are proposed. 103

9.1 Oracle Result of various appliances based on different features 106

9.2 Comparison between trivial and Neural Network Learner 107

10.1 Accuracy of the prediction with two different kind of inputs 121

10.2 Prediction results using different multi-label algorithms. 122

10.3 Appliances identification and prediction, using the LP algorithm. 125

10.4 Real-time appliance usage prediction based on a previous identification. 126

11.1 Prédiction de l'état de charge avec identification préalable, ou avec mesures directes170

Table of Abbreviations and Acronyms

Auc	Area under curve
Amm	Automated Metering Management
Ann	Artificial neural network
Bn	Bayesian Network
Br	Binary relevance
Cpt	Conditional probability table
Ddsc	Discriminative disaggregation sparse coding
Dmc	Dynamic matrix control
Dsm	Demand side management
Dtl	Decision tree learner
Dtm	Decision table mapping
Dtw	Dynamic time warping
Ems	Energy management system
Edf	Electricité de France
Erdf	Electricité Réseau Distribution de France
Has	Home automation system
Hmm	Hidden markov models
Hvac	Home ventilation and air conditioning
Knn	K-nearest neighbors classifier
Lp	Label powerset
Ml-Knn	Multi-label K-nearest neighbors
Mpc	Model predictive control
Nialm	Non intrusive appliance load monitoring
Nilm	Non intrusive load monitoring
Pca	Principal component analysis
Pv	Photovoltaic
Rte	Réseau de transport d'électricité
Svm	Support vector machine
Stlf	Short term load forecasting
Tc	Temporal Correlation

French Extended Abstract

RESUME – Nous abordons dans ces travaux l'identification non intrusive des charges des bâtiments résidentiels ainsi que la prédiction de leur état futur. L'originalité de ces travaux réside dans la méthode utilisée pour obtenir les résultats voulus, à savoir l'analyse statistique des données (algorithmes de classification). Celle-ci se base sur des hypothèses réalistes et restrictives sans pour autant avoir de limitation sur les modèles comportementaux des charges (variations de charges ou modèles) ni besoin de la connaissance des changements d'état des charges. Ainsi, nous sommes en mesure d'identifier et/ou de prédire l'état des charges consommatrices d'énergie (et potentiellement contrôlables) en se basant uniquement sur une phase d'entrainement réduite et des mesures de puissance active agrégée sur un pas de mesure de dix minutes, préservant donc la vie privée des habitants. Dans cette communication, après avoir décrit la méthodologie développée pour classifier les charges et leurs états, ainsi que les connaissances métier fournies aux algorithmes, nous comparons les résultats d'identification pour cinq algorithmes tirés de l'état de l'art et les utilisons comme support d'application à la prédiction. Les algorithmes utilisés se différentient par leur capacité à traiter des problèmes plus ou moins complexe (notamment la prise en compte de relations entre les charges) et se ne révèlent pas tous appropriés à tout type de charge dans le bâtiment résidentiel.

Mots-Clés – Identification non intrusive, prédiction, bâtiment résidentiel, classification multi-étiquette, analyse de données.

11.4 Contexte des travaux

Le contrôle des charges dans l'habitat (principalement résidentiel) est un domaine de recherche en plein développement, profitant des opportunités offertes dès lors que les méthodes d'identifications et de prédiction sont suffisamment précises [Berges 2010, Palensky 2011]. De nombreuses approches sont proposées dans la littérature, basées principalement sur une analyse temporelle des signaux électriques et nécessitant un taux d'échantillonnage important (la seconde ou moins) afin de détecter les variations d'état des charges [Fernandes 2013, Siano 2014]. En effet, les changements d'état des charges se reflètent sur leur consommation ce qui permet de détecter leur présence dans le profil de consommation global du bâtiment. Ces méthodes sont donc basées sur la reconnaissance ou la connaissance à priori des signatures des charges, par analyse plus ou moins complexe des signaux électriques, de leur harmoniques ou des contenus supplémentaires que pourrait mesurer par exemple un compteur (très) intelligent [Figueiredo 2011].

La principale limite de ces méthodes est la nécessité de mesures précises ou de modèles fidèles des différentes charges dans l'habitat afin de calibrer correctement les algorithmes sous peine de voir leur efficacité diminuer. De plus, la période d'entrainement de ces algorithmes peut s'avérer suffisamment longue pour rendre leur utilisation impraticable. Le second défaut, directement lié au premier, est l'intrusion dans l'espace privé que constitue l'habitation résidentielle. En effet, que ce soit lors du calibrage des algorithmes, puis par la quantité d'informations précises recueillies sur le mode de vie des habitants, ces méthodes peuvent présenter un caractère rebutant à des utilisateurs soucieux de conserver un certain contrôle sur la connaissance que pourraient avoir des tiers de leur vie quotidienne.

Pour répondre à ces questions (et tenter d'éviter ces écueils) nous proposons une approche qui ne se base pas sur de l'analyse temporelle des signaux mais sur des outils statistiques d'analyse de donnée [Hart 1992, Laxman 2006, Birt 2012, Tsoumakas 2010]. De plus, nous nous plaçons dans le cas d'une mesure unique d'énergie agrégée sur un pas de temps compris entre dix minutes et une heure [Norford 1996]. Ceci nous donne plusieurs avantages. Tout d'abord, les besoins en mesure sont notablement réduits (que ce soit en volume de donnée à traiter et en complexité technique de la mesure : un compteur classique mesurant la puissance active suffit) [Kolter 2010]. D'autre part, nous ne sommes pas en mesure d'identifier précisément le comportement des habitants, mais uniquement d'isoler l'état des charges de forte puissance (le seuil étant défini lors du paramétrage de l'algorithme), c'est-à-dire celles qui présentent de toute façon un intérêt réel à être contrôlées dans l'habitat [Kalogridis 2010]. Enfin, nous avons également limité la phase d'entrainement de l'algorithme à une période très courte (deux semaines sur des données d'un an, soit moins de 4% de la base de données). Tout ceci est fait en se basant sur des informations simples à obtenir sans mesures pour l'identification (il faut uniquement un marquage du temps lors de l'utilisation des charges dans la maison, sans mesures sur la charge elle-même), et avec une connaissance de l'état des charges au fil du temps dans le cas de la prédiction [Kadous 2002].

Ceci nous amène à une méthode quasi-non intrusive d'identification et/ou de prédiction de l'état des charges dans l'habitat avec beaucoup moins de contraintes que des méthodes temporelles [Hart 1992, Dong 2012, Zeifman 2011]. En plus du respect de la vie privée, cette méthode va s'adapter à n'importe quel type de charge (c'est-à-dire quel que soit sa courbe de charge, ou son modèle électrique) [Labeeuw 2013]. Le modèle des charges n'ayant pas à être connu, nous n'avons pas besoin d'une mesure au niveau de la charge pour le construire. A chaque ajout ou changement de charge, ou alors sur une base régulière pour prendre en compte le vieillissement des charges et la dégradation de leur façon de consommer de l'énergie, il suffit de relancer une courte période d'entrainement de l'algorithme afin de le rendre fonctionnel sur la nouvelle configuration de la maison [Onoda 2000].

Les algorithmes développés pour l'identification de l'état des charges peuvent également être utilisés pour prédire leur état futur. La différence est qu'il faut cette fois-ci une mesure au niveau des charges donnant au minimum une connaissance de leur état. Cette connaissance peut être apportée par les mêmes algorithmes, par identification, en utilisant ainsi uniquement la courbe de charge globale de la maison. Dans ce cas, l'efficacité globale de la méthode est moins bonne.

Dans cette communication, nous présentons les résultats d'identification des charges selon plusieurs algorithmes d'analyse et de classification de données, fournissant ainsi une comparaison à l'état de l'art. Puis nous appliquons les mêmes algorithmes pour réaliser de la prédiction de charge afin de montrer la viabilité de la méthode.

11.4.1 Les données

Dans le cadre du projet ANR SUPERBAT, nous disposons pour nos travaux d'une base de donnée, nommée IRISE, tirée d'un projet européen nommé REMODECE. Celle-ci contient le relevé de consommation, pour cent maisons et sur un an, de la quasi-totalité des charges sur un pas de temps de dix minutes. Les relevés se présentent sous la forme d'agrégation de puissance sur un pas de temps de dix minutes, donc finalement comme une représentation de l'énergie consommée sur les dix dernières minutes. Des données de températures et d'autres informations complètent cette base de données (comme le nombre d'habitant par exemple).

Les algorithmes d'identification et de prédiction présentés dans cet article sont basés sur des algorithmes de *classification*. Ceux-ci procèdent par apprentissage. Nous coupons donc la base de données en deux parties, la première pour l'apprentissage et la seconde pour la validation. La répartition dans notre cas est de 4 % pour l'apprentissage et de 96 % pour la validation. Ce choix non typique (certains algorithmes nécessitent un apprentissage sur parfois plus de 50 % de la base de donnée) est une contrainte supplémentaire que nous imposants à la méthode en vue de la rendre plus applicable à une situation réelle. En chiffres concrets, pour fournir des informations sur les charges d'un bâtiment pendant un an, l'algorithme ne nécessite qu'une période d'apprentissage de deux semaines sans mesures directe sur les charges (juste une prise de note de l'heure lors du changement d'état des charges).

Le choix de pas de temps d'une dizaine de minutes ou de l'heure est un second choix restrictif imposé dans cette étude. Celui-ci nous permet de rester dans des temps représentatifs de remontée de mesures aux réseaux sur les prototypes de compteurs intelligents (les mesures en local pouvant se faire à un échantillonnage plus fin) et de ne pas porter atteinte à la vie privée des habitants.

11.5 Classification des charges

11.5.1 Sélection des charges

L'objectif de ce travail est d'isoler l'état d'une charge consommant suffisamment d'énergie pour présenter un intérêt à être contrôlée dans une maison, à partir de la seule connaissance de l'évolution de la consommation globale (en puissance active) de cette maison.

Le seuil de discrimination entre les charges « intéressantes » et les autres se fait sur leur niveau de consommation d'énergie moyenne sur un pas de temps de dix minutes. Ce seuil est défini d'une part pour faciliter l'identification des charges « utiles ». En effet, il est inutile de brouiller la courbe de consommation des charges contrôlables de forte puissance avec du bruit de mesure ou des charges consommant trop peu pour être faciles à distinguer les unes des autres. Les exemples classiques de charges de ce type sont les lampes, les chargeurs de téléphones portables, etc. Ce seuil est défini d'autre part pour mettre naturellement de côté les charges qui, même si elles étaient contrôlables, n'apporteraient pas une flexibilité suffisante sur la courbe de consommation globale pour justifier la mise en place de toute une mécanique de classification afin de détecter automatiquement leur état, ou plus tard de le prédire et le contrôler plus finement.

Ce seuil peut être choisi arbitrairement par les habitants en fonction de leur moyenne globale de consommation, ou choisi de façon plus systématique par des méthodes de groupement statistique (« clustering ») basé sur un historique de consommation initial. Contrairement à la méthode que nous présentons ici, le groupement automatique des charges en deux catégories (les charges de puissance moyenne suffisante et les autres) nécessite la connaissance de la mesure de consommation de chaque charge prise séparément. Bien que nous ayons à notre disposition les mesures individuelles des charges dans la base de donnée IRISE, nous avons choisi de rester sur un critère de seuil manuel, dans la mesure où notre méthode est forte de ne pas nécessiter ces mesures. Le seuil limite en deçà duquel les charges ne sont pas considérées a été défini sur la base de la consommation moyenne d'une télévision telle que mesurée dans les maisons de la base IRISE utilisées dans ces travaux.

11.5.2 Mesures et classification

L'entrée principale de l'algorithme de classification est la puissance consommée par la maison complète agrégée sur un pas de temps qui peut varier de dix minutes à une heure. Cette énergie est relevée au niveau du compteur. Lors de la phase d'apprentissage, en plus de la consommation globale et du temps, les états des charges (ON-OFF) sont fournis à l'algorithme. Lors de la phase d'utilisation, l'algorithme n'a plus à sa disposition que le temps et la consommation globale de la maison. Il donne à ce moment-là une proposition de l'état des charges, accompagné d'indices de confiances.

Les charges considérées sont soit allumées/utilisées (ON) soit éteintes/non utilisées (OFF). Il n'y a pas de mode de veille.

Deux remarques importantes sont à faire ici. Tout d'abord cette simplification du problème pour une meilleure lisibilité n'exclue pas de classer les charge sur plus d'étiquettes que cela (qui seront alors dans notre cas les niveaux d'énergie consommés sur les pas de temps considérés et non plus des états ON ou OFF). Dans le cas d'une discrétisation plus fine que juste l'état de la charge, le choix de considérer de la puissance ou de l'énergie vient directement de la mesure qui est faite sur les charges réelles, dans notre cas de l'énergie (puissance agrégée sur un pas de temps de dix minutes). D'autre part, afin d'aller plus loin que la seule connaissance de l'état de la charge, il faut impérativement une mesure de la consommation de la charge lors de la phase d'apprentissage de l'algorithme (en puissance ou énergie par pas de temps). Ceci, tout à fait possible en phase de recherche, devient plus difficile une fois que ce système est à déployer en situation réelle, ou lors d'une mise à jour du nombre de charges et de leur état. Il faut en effet installer des appareils de mesures sur chaque charge pour une durée pouvant aller jusqu'à plusieurs semaines pour de l'identification et pour la durée d'usage pour de la prédiction.

Il est à noter que cette mesure pourrait à terme être centralisée par une « énergie box », ou un compteur de type Linky, comme présenté Figure 11.11. A ce stade, l'identification des charges telle que nous la présentons devient obsolète dans la mesure où les informations sur l'état des charges sonst en permanence relevées ou du moins potentiellement connues par le système de management énergétique du bâtiment par une simple requête informatique. Ceci étant dit, le déploiement massif « d'énergie box » en plus de compteurs communicants (qui ne seraient pas suffisants) ne semble pas être encore une réalité à l'heure actuelle et dans tous les cas, la méthode présentée ici reste utilisable sur son aspect prédictif.

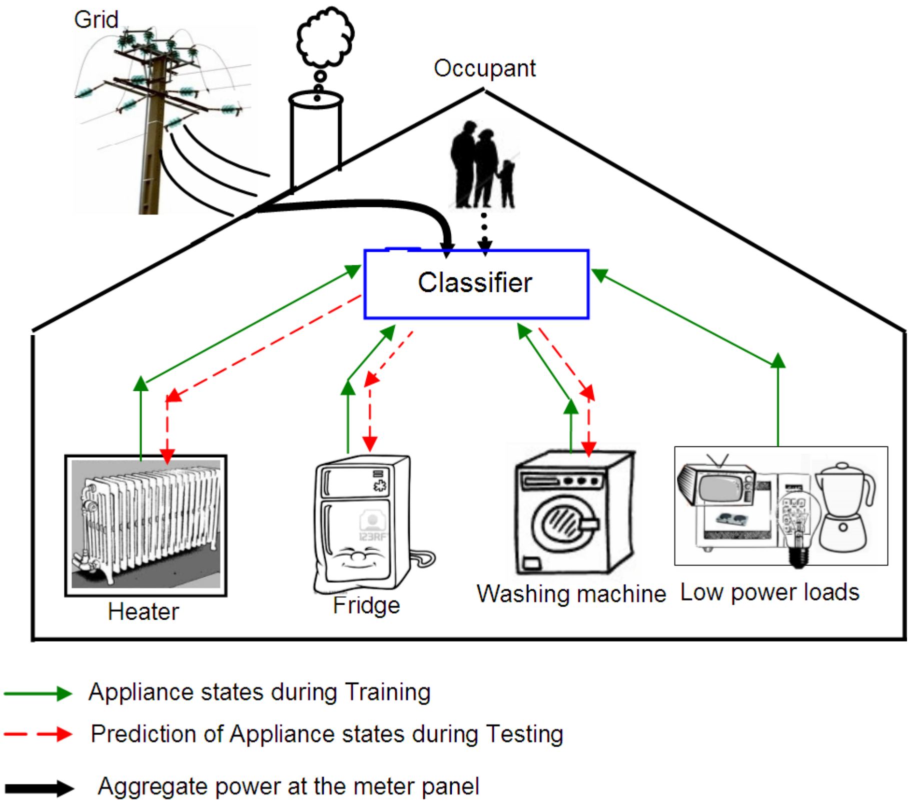


FIGURE 11.11 – Architecture de la classification par l'algorithme d'apprentissage

11.5.3 Principe de fonctionnement

La méthode utilisée dans ces travaux se base sur la classification de série de données temporelles selon une approche par fenêtres de mesures successives. Il s'agit de classer les charges en fonction de leur état (ON ou OFF) ce qui permettra ensuite de les identifier en se basant uniquement sur le profil de charge global de la maison. Il s'agit donc d'une classification multi-classe et multi-étiquettes (ou multi-état). Chaque charge de la maison prise en compte dans l'algorithme est définie comme une classe qui peut comporter deux étiquettes, ON ou OFF. Le travail de classification se déroule en plusieurs étapes, qui sont décrites succinctement ci-dessous.

1. Pré-traitement des données.
2. Calcul des attributs/propriétés pour chaque charge.
3. Apprentissage de la classification multi-étiquette.
4. Classification des charges.
5. Evaluation de la pertinence de la classification

Sans rentrer dans les détails informatiques de ces étapes, le principe de fonctionnement de l'algorithme est décrit ci-dessous.

11.5.4 Les étapes clés

11.5.4.1 Découpage de la série temporelle initiale

La classification s'opère par analyse comparative de l'évolution de séries temporelles. Il faut donc découper la série initiale S construite à partir de l'énergie consommée par le bâtiment agrégée sur des pas de dix minutes à une heure en un ensemble de séries S_i, de longueur n, extraites de S par découpage en fenêtres successives, comme présenté, eq. (11.21). La taille des fenêtres peut varier (en nombre d'unité ou au niveau du centrage par rapport au pas de temps analysé à chaque instance) [Basu 2012c].

$$S = \{s_1, s_2, ...s_L\} \rightarrow \begin{cases} S_1 & = \{s_1, s_2, ..., s_n\} \\ S_2 & = \{s_2, s_3, ...s_{n+1}\} \\ & \vdots \\ S_N & = \{s_{L-n-1}, s_{L-n-2}, ...s_L\} \end{cases} \tag{11.21}$$

A partir d'une série initiale S de longueur L (directement obtenue par le choix de discrétisation de la base de donnée d'un an à notre disposition) nous extrayons donc $N = (L - n)$ séries par découpage en fenêtres successives, séparées d'une unité de pas de temps. Cette unité de pas de temps a été prise de dix minutes et d'une heure pour les résultats de cet article. La longueur de la sous-séries S_i, notée n est un choix qui aura une influence sur l'efficacité d'identification de l'algorithme. Il s'agit d'une valeur à déterminer au cas par cas, car il faut être capable de prendre en compte une plage assez large pour capturer des variations d'état de charge, mais pas trop pour fournir à l'algorithme le plus de séries possibles pour faciliter son travail de classification. Après plusieurs essais, avons choisi de travailler systématiquement dans nos travaux avec une fenêtre de découpage de dix unités, l'unité étant le pas de temps d'échantillonnage de la base de données initiale.

Lors de la phase d'apprentissage, l'algorithme parcours la série initiale S par étapes successives (les S_i) en suivant l'information temporelle qui lui est fournie. Il prend donc en compte par construction la succession temporelle des évènements. A chaque pas de temps t_i, l'algorithme choisi les sous-série correspondantes S_i. Il est possible de prendre du premier élément au dernier de cette série S_i de longueur n comme référence pour le placement du temps. Après avoir vérifié que la sensibilité de la classification des charges est faible par rapport au choix de la position de l'élément de référence dans cette sous-série, nous avons choisi de centrer la sous-série S_i sur le pas de temps t_i [Basu 2012c].

11.5.4.2 Caractéristiques supplémentaires

Pour chacune des fenêtres successives considérées, nous calculons des informations supplémentaires caractéristiques que l'algorithme est susceptible de lier à l'état des charges (qu'il connait lors de la phase d'entrainement). Suivant les algorithmes, il est également en mesure de lier les états des charges entre eux. Par exemple, il y a des chances que le sèche-linge soit utilisé après la machine à laver. Cette liaison représente une information qui sortira avec plus d'importance dans les arbres de décisions de classification des différents algorithmes, s'ils sont capables de les lier.

Lors de la phase d'apprentissage, il faut donc fournir à l'algorithme, en plus de l'état des charges, du temps et de la consommation globale le plus d'informations susceptibles d'améliorer sa catégorisation des charges. Ainsi, un point clé du travail est de trouver les bonnes propriétés avec lesquels alimenter l'algorithme de classification. En effet, fournir des informations non pertinentes ne peut que dégrader l'efficacité d'identification et de prédiction. Ces informations sont calculées sur chacune des fenêtres de calculs présentées ci-dessus.

Un exemple d'information analysée sur une fenêtre de calcul est la variation en énergie sur les pas de temps suivants et précédents le pas de temps de l'itération considérée, comme présenté sur la Figure 11.12 où nous pouvons également positionner les maximum et minimum locaux de la fenêtre d'observation.

Les informations complémentaires apportées à l'algorithme de classification représentent la valeur ajoutée de la connaissance des charges que ne peut deviner un algorithme. En plus de la position du maximum et du minimum de consommation dans la fenêtre considérée, nous ajoutons les dérivées première et seconde de cette consommation globale, des analyses statistiques sur la consommation (variation entre chaque pas de temps, moyenne et écart type). De plus, des informations temporelles supplémentaires sont données à l'algorithme : l'heure de la journée (de 0 à 23) et le jour de la semaine. Ceci permet de relever des schémas temporels sur les courbes de charge sans directement les connaître. Avec cette information, l'algorithme sera en mesure de classer les charges correspondantes lorsqu'il les reconnaitra dans la consommation globale en phase de validation, en usage journalier ou hebdomadaire.

Chaque fenêtre d'observation successive est analysée par l'algorithme afin qu'il se construire une base de connaissance en phase d'entrainement qu'il utilisera en phase de validation en ayant cette fois-ci a

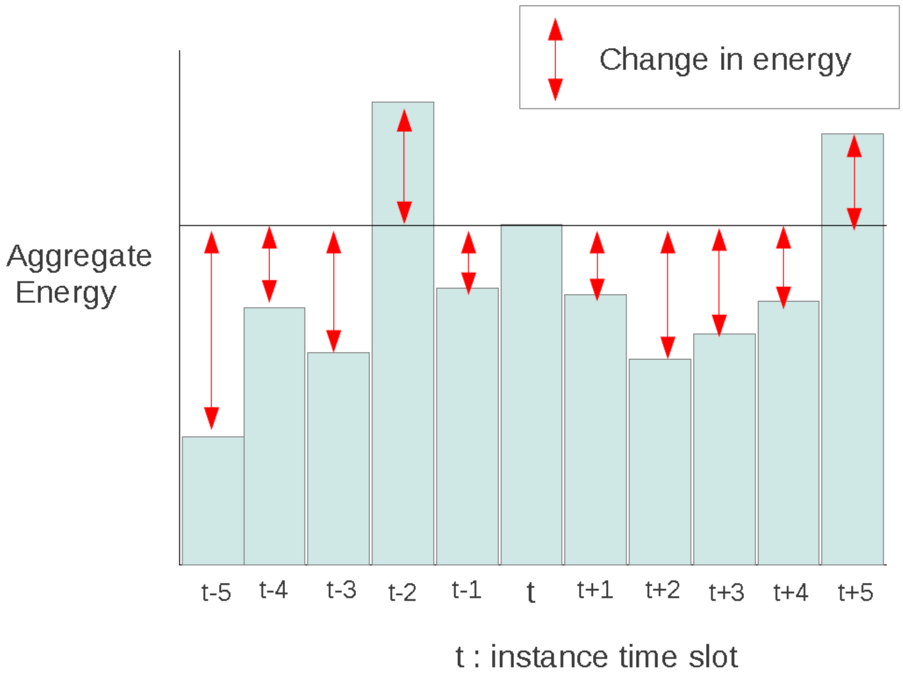


Figure 11.12 – Fenêtre de calcul et caractéristiques

sa disposition uniquement la consommation globale de la maison ainsi que des informations temporelle (heure du jour, journée de la semaine). Le synoptique de fonctionnement des algorithmes de classification est présenté Figure 11.13.

11.5.5 Les algorithmes de classification

Nous avons implémenté cinq algorithmes de classification représentatifs de l'état de l'art dans le domaine. Ces algorithmes se distinguent par leur construction qui les rends naturellement ou artificiellement multi-classe. Dans le second cas, ils ne sont pas capables de lier des évènements entre classes (donc entre charges pour nous). Dans ce cas, il n'est pas possible par exemple de lier l'usage du sèche-linge avec celui de la machine à laver afin d'augment l'efficacité de classification.

Leur nom et les caractéristiques principales des algorithmes sont résumés ci-dessous.

LP1 : *Label powerset problem transformation*, utilisant un algorithme d'arbre de décision.

LP2 : *Label powerset problem transformation*, utilisant une classification *support vector machine*.

BR1 : *Binary relevance problem transformation*, utilisant un algorithme d'arbre de décision.

BR2 : *Binary relevance problem transformation*, utilisant une classification *support vector machine*.

MLkNN : *Multi-label k Nearest Neighbors*, avec K=7.

Une description succincte des algorithmes est proposée ci-dessous [Basu 2013a]. Pour rappel, dans notre cas les classes sont les charges et les étiquettes (« label » en anglais) sont leurs états (On ou Off).

Binary Relevance, BR : Il s'agit d'une méthode de transformation de problème qui effectue un apprentissage séparé par classe sur uniquement deux étiquettes (dans notre cas On et Off) d'où le nom de binaire. L'algorithme BR effectue une transformation des séries initiales (par classe) en une seule série mono-classe contenant toutes les données de toutes les classes. Ensuite, l'algorithme extrait autant de tableaux qu'il y a d'étiquettes, chacun d'eux regroupant tous les attributs liés à l'étiquette.

Label Powerset, LP : L'algorithme LP considère l'ensemble des étiquettes de l'ensemble des classes comme une classe unique ayant une seule étiquette. Contrairement à l'algorithme BR, l'algorithme LP

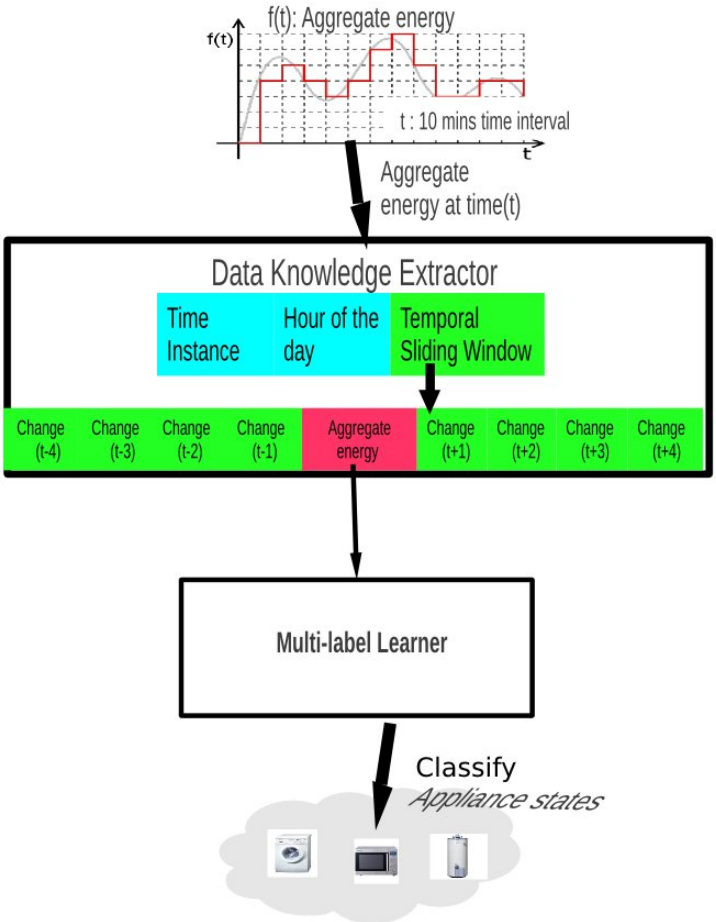


FIGURE 11.13 – Synoptique d'identification des charges

effectue son apprentissage sur un seul classificateur constitué de l'ensemble des données. L'intérêt principal est qu'il est donc possible de prendre en compte la relation entre les charges, ce que ne peut pas faire le BR. La contrepartie est le temps de calcul.

Classification : Pour ces deux algorithmes, il faut faire intervenir un second algorithme qui va classer l'apparition d'évènements après transformation des données. Dans notre cas, ce sont les algorithmes « Decision Tree Learning » (DTL) et « Support Vector Machine » (SVM).

DTL est un moyen de visualiser les relations entre les informations à disposition et l'identification des charges sous la forme d'un arbre de décisions. Etant très visuel, il est systématiquement testé pour valider ses résultats.

SVM permet de grouper les charges entres elles et donc de les identifier à partir des informations mises en forme (en terme de classes et d'étiquettes) par les algorithmes LP et BR.

Multi-label k Nearest Neighbors : Cet algorithme calcule directement des distances entre séries temporelles dans l'espace qui nous convient (en fonction des valeurs d'étiquettes) permettant directement une classification et donc l'extraction d'informations sous forme d'état des charges. L'efficacité de l'identification se fera suivant le choix des paramètres constitutifs de l'algorithme. Nous avons choisi ici $k = 7$.

11.5.6 Performances de l'identification

Les résultats de classification sont exprimés en calculant un score d'identification des charges appelé *F-measure*. Celui-ci est définit comme une moyenne de deux autres indicateurs :

Précision : Pourcentage des états positifs (ON) correctement identifiés.

Recal : Rapport entre le nombre d'états positifs (ON) et le nombre total d'états positifs correctement prédits.

11.5.7 Prédiction

L'usage des algorithmes de classification pour effectuer de la prédiction est similaire à celui qui en est fait pour l'identification. Le principe de fonctionnement est proposé Fig. 11.14.

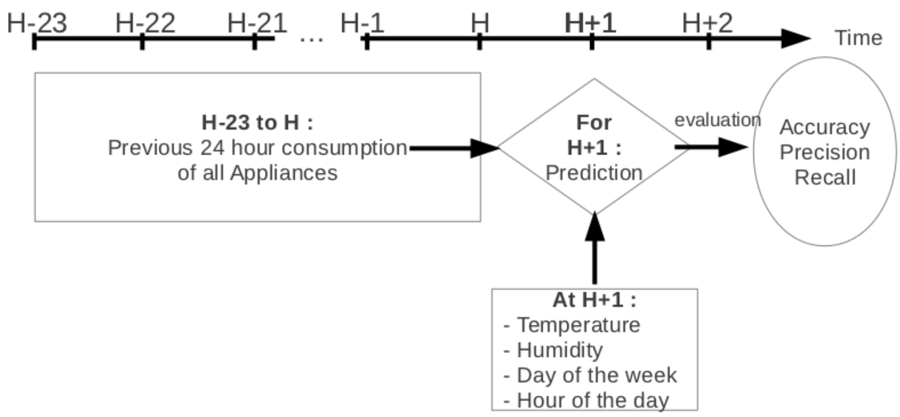


FIGURE 11.14 – Principe de fonctionnement de l'algorithme prédictif

Les algorithmes de classification sont utilisés ici pour proposer un état futur des charges en fonction de la connaissance d'un historique qui grandit à chaque instant et une phase entrainement initiale. Cela revient à identifier l'état d'une charge sans connaitre encore la valeur de la consommation globale de la maison, c'est-à-dire en se basant sur des schémas similaires apparaissant dans l'historique de mesures.

11.6 Résultats

Nous avons classé nos résultats en fonction de catégories de scénarios de calcul (et plus tard de topologie des maisons). Dans un premier temps, nous distinguons les cas où il y a :

1. Peu de charges de forte consommation, identifiées séparément.
2. Peu de charges de forte consommation dont certaines sont identifiées ensemble.
3. Beaucoup de charges de forte consommation avec des doubles possibles (par exemple deux télévisions).

Afin d'estimer la robustesse de l'algorithme d'identification à des configurations différentes sur le nombre de charges et leurs caractéristiques, nous avons mené notre étude sur les trois catégories de maisons ci-dessus. Cependant, les résultats présentés ici ne concernent que la première catégorie de maison.

Les deux autres catégories de maisons sont plus difficiles à traiter pour les algorithmes de classification qui donnent des résultats moins bons. La différence avec les résultats obtenus sur la première catégorie permet de les caractériser. N'étant pas l'objet de cet article, ce ne sera pas traité ici.

11.6.1 Identification

Nous proposons Figures 11.15 et 11.16 la comparaison des cinq algorithmes dont la description est donnée en Section 11.5.5 en fonction de leur performance unifiée d'identification d'état des charges de fortes puissance dans une maison donnée. L'indicateur utilisé pour mesurer la performance est *F-measure*, présenté en Section 11.5.6. La Figure 11.15 est calculée pour un taux d'échantillonnage de dix minutes alors que la Figure 11.16 pour un taux d'échantillonnage d'une heure.

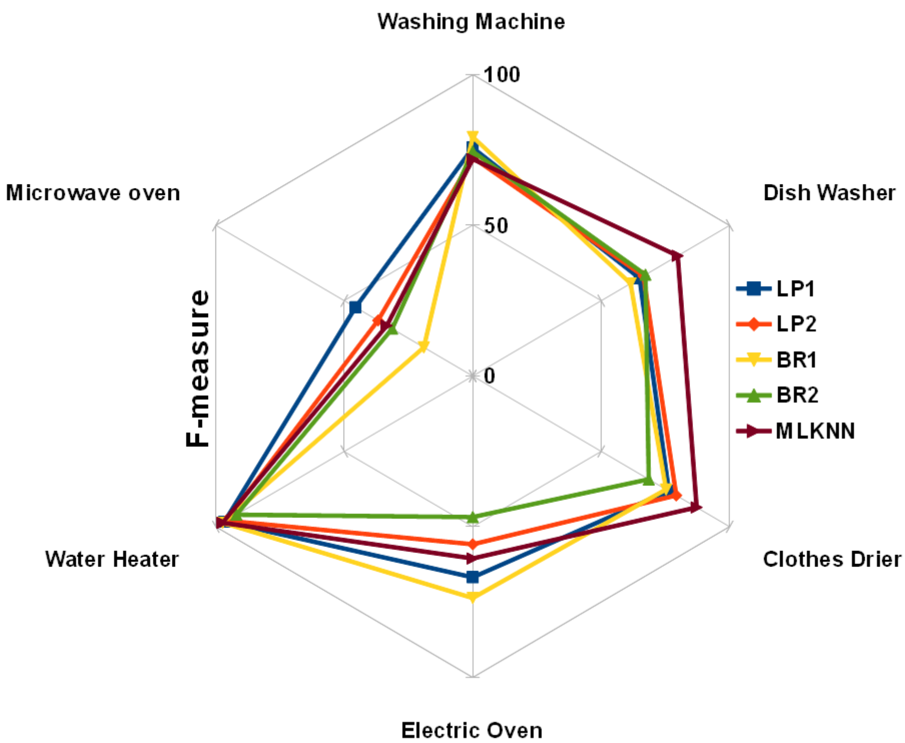


FIGURE 11.15 – Identification d'état des charges (pas de dix minutes)

Nous pouvons constater sur ces figures que l'identification est sensible au choix des algorithmes (notamment ceux qui sont multi-étiquette par défaut ou non) qui vont de fait présenter des résultats variables en fonction des charges considérées. Cependant, les différences ne sont pas aussi grandes entre algorithmes qu'entre pas de temps d'échantillonnage.

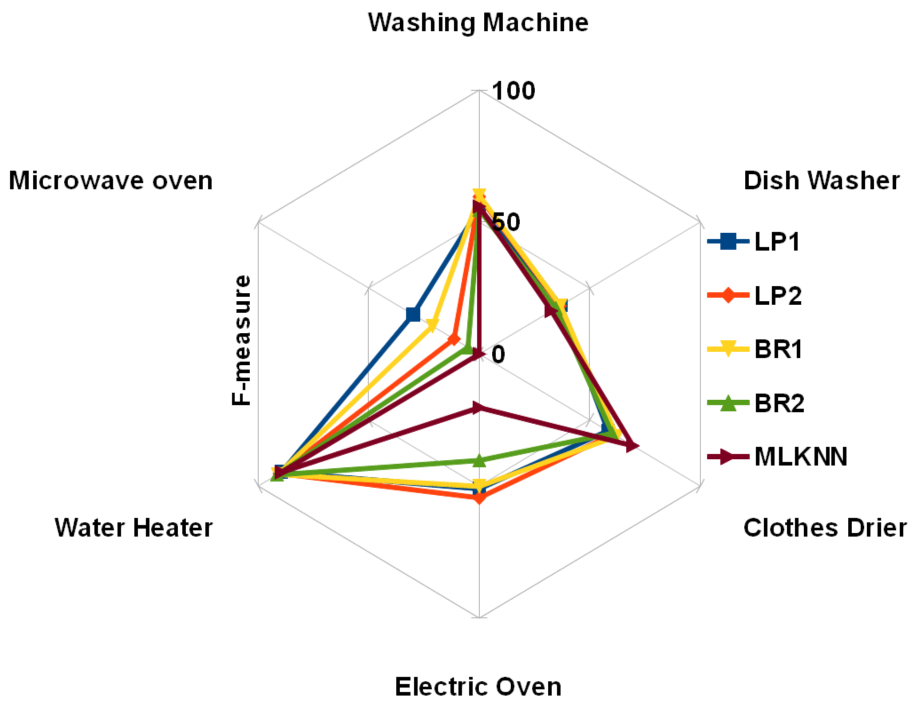


FIGURE 11.16 – Identification d'état des charges (pas d'une heure)

La charge la plus difficile à identifier est le four à micro-ondes. Pour les autres charges, l'algorithme MLkNN est quasiment à chaque fois le plus performant, secondé, pour les charges qui présentent un lien entres elles (sèche-linge et machine à laver) par les algorithmes multi-classe (type LP).

L'augmentation du temps d'échantillonnage a un impact significatif sur les résultats d'identification, mais finalement moins important que ne pourrait le laisser penser le passage d'un échantillonnage de dix minutes à une heure. D'autre part, l'identification se fait avec des résultats très corrects pour la plupart des charges ayant vocation à être contrôlées, même avec une mesure toutes les dix minutes. C'est le résultat notable de cette étude en vue d'une implémentation dans une première génération de compteurs communicants.

11.6.2 Prédiction

Les résultats de prédiction sont exprimés en termes de confiance, comprise entre zéro (aucune confiance dans la prédiction) et un (confiance totale dans la prédiction). Nous proposons également une représentation de la fréquence d'usage en fonction de l'heure de la journée.

Figure 11.17 propose les résultats de prédiction comparés aux mesures réelles, présentés sous la forme d'un diagramme fréquentiel horaire. Une prédiction est considérée comme réussie lorsque la charge considérée est dans le même état pendant les 10 minutes du pas de temps. En plus de renvoyer beaucoup d'informations aux gestionnaires réseau, ce type de graphique représente également des informations potentiellement utiles à titre informatif pour les occupants (en vue d'améliorer leur efficacité énergétique).

Figure 11.18 propose les résultats de confiance pour la prédiction de l'usage de la télévision dans une des maisons de la base de données IRISE. Le seuil considéré pour tracer ce graphique est sévère (ON ou OFF sur 10 minutes) ce qui implique un certain nombre de ratés sur une semaine, mais reste une indication fiable sur la plage horaire d'observation.

Nous pouvons constater sur ces figures que la prédiction est viable pour les charges de forte puissance typiques que nous présentons. Ceci n'est bien sûr pas toujours le cas, notamment dans les maisons possédant plusieurs fois la même charge (deux téléviseurs par exemple). Ce point est négatif du point de vue d'un management local des charges avec comme bénéficiaire l'utilisateur (par exemple pour réduire sa facture d'électricité) mais reste pertinent du point de vue du gestionnaire réseau, car les applications visées sont plus rares, et peu dupliquées dans l'habitat (machine à laver, chauffe-eau, etc.).

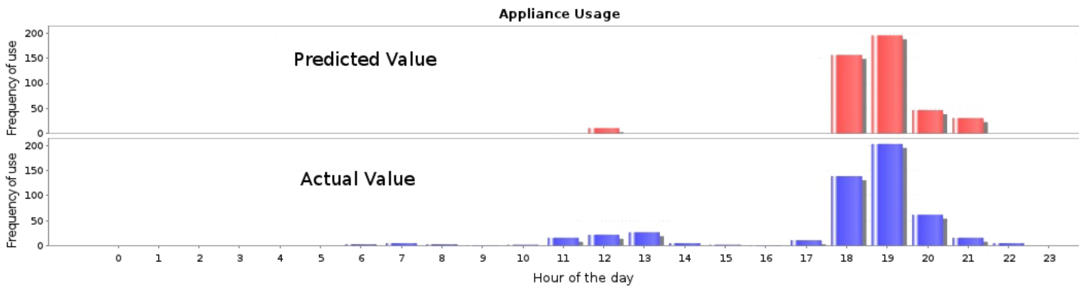


FIGURE 11.17 – Diagramme de fréquence horaire, four électrique (plaques et four)

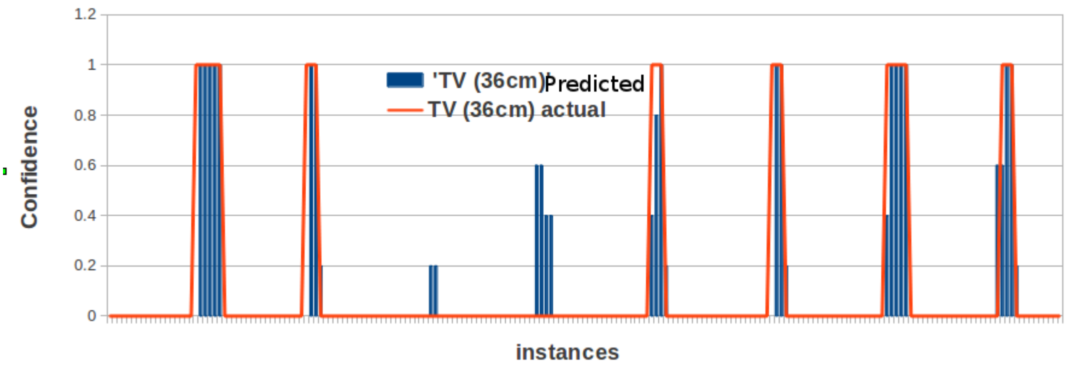


FIGURE 11.18 – Confiance de prédiction d'utilisation d'un téléviseur sur une semaine

11.6.3 Identification puis prédiction

Comme dit précédemment, l'algorithme de classification peut être utilisé pour la prédiction. La contrepartie est qu'il faut lui fournir l'état des charges les instants précédent l'instant à prédire. Deux solutions sont possibles. Soit l'état des charges est mesuré et il est utilisé directement comme entrée pour l'algorithme. Soit seule la courbe de charge est disponible, il faut donc identifier l'état des charges avec l'algorithme de classification, dont le résultat sera donné à l'étage de prédiction (basé sur les mêmes principes de classification). Cette deuxième solution a une précision plus faible, dans la mesure où les erreurs cumulées lors de l'identification impactent directement la prédiction. Son principe est proposé Fig. 11.19.

L'usage successif de l'identification et de la prédiction, ou la prédiction seule sont illustrés dans le tableau 11.1. Il y est proposé la comparaison de la prédiction de l'état des charges dans le cas d'une connaissance directe de l'état des charges ou d'une identification préalable. La différence entre les deux prédictions est directement liée aux erreurs cumulées lors de l'étape d'identification, faibles, mais inévitables.

Les résultats présentés dans le tableau 11.1 montrent bien la dégradation des résultats en associant deux étages de classification sans mesures réelles. Cependant, ces résultats ne sont pas non plus mauvais. De plus, ils sont variables selon les charges considérées, donc présentent tout de même un intérêt à l'usage.

Dans le cas où une mesure directe est disponible, les résultats de prédiction sont suffisant pour une utilisation réelle, ce qui valide l'intérêt de la méthode.

11.6.4 Limites de la méthode

La robustesse de la méthode lorsque le nombre de charges et leur caractéristiques évoluent dans le temps par rapport à la configuration initiale utilisée dans l'apprentissage n'a pas été étudiée. En principe, il est très simple de lancer à nouveau une phase d'apprentissage. D'autre part, ces méthodes ne sont pas limitées en nombre ni type de charges, car les algorithmes de classification détermine seul les catégories de charges et leurs étiquettes associées.

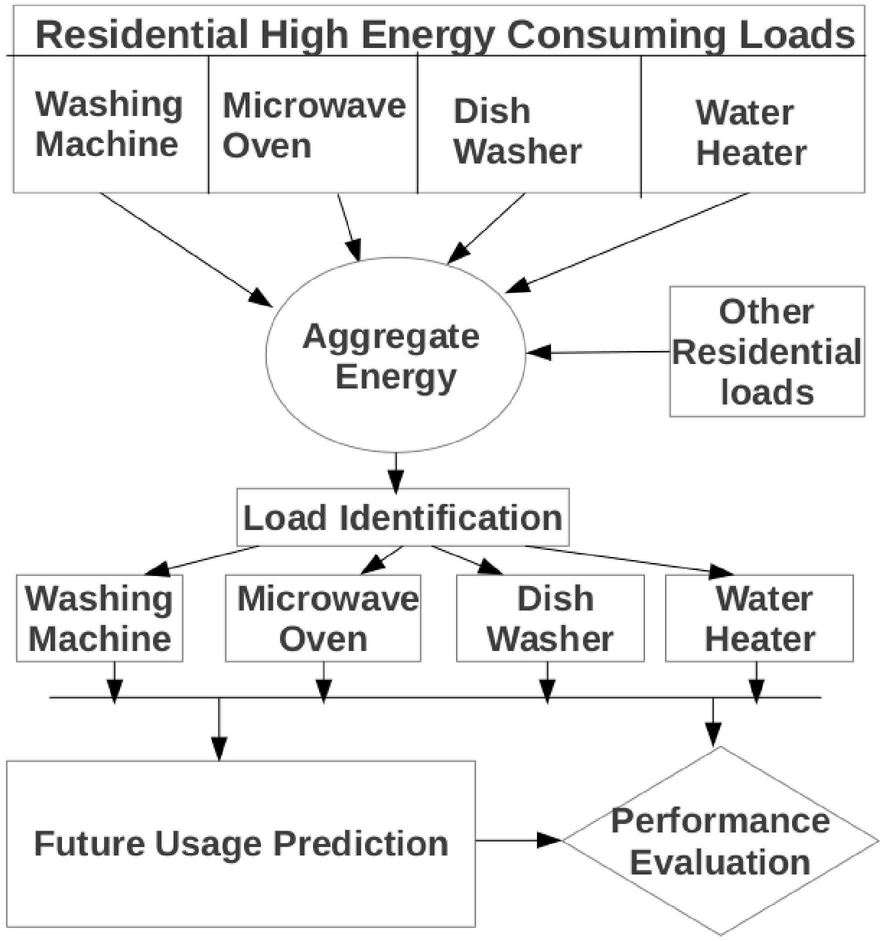


FIGURE 11.19 – Principe d'association des algorithmes d'identification et de prédiction

TABLE 11.1 – Prédiction de l'état de charge avec identification préalable, ou avec mesures directes

Charges	Algorithmes	Basé sur l'identification (Smart Meter)			Pas d'identification (mesure directe)		
		Accuracy	Precision	Recall	Accuracy	Precision	Recall
Machine à laver	LP	95.11	66.66	18.66	96.58	90.74	64.42
	BR	95.13	60.51	28.22	96.61	90.00	65.57
Four à micro-onde	LP	88.22	13.33	1.41	90.47	32.83	2.75
	BR	88.27	0	0	90.40	35.92	4.62
Chauffe-eau	LP	95.71	83.42	81.68	98.73	96.29	93.29
	BR	95.96	86.16	80.33	98.73	96.29	93.29
Lave-vaisselle	LP	95.94	0	0	98.96	83.67	33.60
	BR	95.94	0	0	99.00	86.00	35.24

Virtuellement, la robustesse est donc très grande. D'autant plus (pour la prédiction) qu'elle est mise à jour à chaque pas de temps. Cependant, elle reste à être évaluée.

11.7 Conclusion

Du point de vue informatique, l'originalité de ce travail est basé sur le calcul et la prise en compte d'informations spécifiques tirées de notre expérience dans le domaine du Génie Electrique fournies aux algorithmes plus classiquement utilisés dans le domaine du « text-mining » par exemple. Du point de vue plus technique, nous mettons en avant une méthodologie qui a pour avantage la prise en compte de limitations pratiques sans pour autant garder d'hypothèses de fonctionnement restrictives sur les charges pour l'aide à l'identification. Un exemple d'hypothèse couramment utilisée dans les analyses temporelles et qui n'est pas appliquée ici est le fait que deux charges ne peuvent pas changer d'état en même temps.

Avec des informations restreintes et volontairement non-intrusive, nous sommes en mesure d'identifier les charges dans une habitation résidentielle sans avoir besoin de mesurer leur variations de consommation (pas d'identification des transitions d'état requises). Les charges considérées sont celles qui consomment le plus d'énergie dans la maison, et à l'intérieur de cette catégorie, celles qui ont le potentiel d'être contrôlables, que ce soit localement ou à distance.

Nous avons comparé plusieurs algorithmes de l'état de l'art nous permettant d'atteindre ces résultats, qui sont actuellement en cours d'utilisation dans des optimisations réactive locales du triptyque bâtiment – panneaux photovoltaïques – batteries de véhicules électriques sous critères technico-économiques, ainsi que dans une perspective plus durable (prise en compte supplémentaire d'impacts environnementaux). Ces travaux menés à une première échelle limitée sont également développés dans un environnement d'agrégation de charges (par exemple au niveau d'un quartier) à une échelle plus générale cette fois-ci en vue d'estimer un niveau de flexibilité que peut attendre un gestionnaire réseau d'un « quartier intelligent ».

D'autre part, nous avons également appliquée cette méthode de classification pour *prédire* l'état des charges (On ou Off), qui peut être lui-même généralisé à la prédiction de niveaux d'énergie [Basu 2013a]. L'identification et la prédiction que nous proposons ont une base méthodologique similaire, mais une utilisation différente des structures algorithmiques développées (notamment le traitement et l'utilisation des données).

L'identification et la prédiction des charges associées ensemble ont deux intérêts. Tout d'abord, elles permettent de remonter des informations décomposées par charge à l'habitant sur sa consommation, sans appareillage de mesure sophistiqué ni distribué. Il est ainsi possible d'analyser et adapter sa consommation ou détecter qu'un appareil est déficient. D'autre part, dans le cadre d'un contrôle distant, le gestionnaire du réseau de distribution pourra (suivant les évolutions futures de ses capacités de gestion) ajuster de façon plus fine la consommation à la production locale et minimiser ainsi les variations d'échanges de puissance entre différentes sections du réseau par un contrôle approprié des sources locales (qu'elles soient de production ou de consommation). Dans ce cas, l'identification des charges des bâtiments pourra se faire sur des agrégations de maisons réalisées selon un groupement préalable (statistique par exemple) basé sur des profils de consommation et d'autres informations (nombres d'habitants, région, températures, etc.)

Bibliographie

[Abras 2006] S. Abras, S. Ploix, S. Pesty et M. Jacomino. Informatics in control, automation and robotics iii, volume 15 of *Lecture Notes in Electrical Engineering*, chapitre A multi-agent home automation system for power management. Springer Berlin Heidelberg, 2006. Andrade Cetto, J. ; Ferrier, J.-L. ; Pereira, J.M.C.D. ; Filipe, J. (Referenced on page 19.)

[Abras 2008] S. Abras, S. Pesty, S. Ploix et M. Jacomino. *An Anticipation Mechanism for Power Management in a Smart Home using Multi-Agent Systems*. In Information and Communication Technologies : From Theory to Applications, 2008. ICTTA 2008. 3rd International Conference on, pages 1–6, april 2008. (Referenced on page 15.)

[Abras 2010] S. Abras, S. Ploix, S. Pesty et M. Jacomino. *Advantages of MAS for the resolution of a power management problem in smart homes*. In 8th International Conference on Practical Applications of Agents and Multi-Agent Systems, PAAMS'2010, Salamanca, Spain, 26-28 April 2010. Springer Verlag. (Referenced on page 15.)

[Arghira 2012] N. Arghira, L. Hawarah, S. Ploix et M. Jacomino. *Prediction of appliances energy use in smart homes*. Elsevier Energy, 2012. (Referenced on page 116.)

[Badreddine 2013] R.M. Badreddine. *Gestion Energetique Optimisée pour un Bâtiment Intelligent Multi-sources Multi-charges : différents principes de validation*. PhD thesis, Université de Grenoble, 2013. (Referenced on page 132.)

[Bakirtzis 1996] A.G. Bakirtzis, V. Petridis, S.J. Kiartzis, M.C. Alexiadis et A.H. Maissis. *A neural network short term load forecasting model for the Greek power system*. Power Systems, IEEE Transactions on, vol. 11, no. 2, pages 858–863, may 1996. (Referenced on pages 18 et 91.)

[Baronti 2007] P. Baronti, P. Pillai, V.W.C. Chook, S. Chessa, A. Gotta et Y.F. Hu. *Wireless sensor networks : A survey on the state of the art and the 802.15. 4 and ZigBee standards*. Computer communications, vol. 30, no. 7, pages 1655–1695, 2007. (Referenced on page 21.)

[Basu 2011] K. Basu, M. Guillame-Bert, H. Joumaa, S. Ploix et J. Crowley. *Predicting Home Service Demands From Appliance Usage Data*. In International Conference on Information and Communication Technologies and Applications (ICTA), 2011. (Referenced on page 3.)

[Basu 2012a] K. Basu, V. Debusschere et S. Bacha. *Appliance usage prediction using a time series based classification approach*. In Industrial Electronics (IECON), IEEE Conference on, pages 1217–1222, October 2012. (Referenced on pages 13, 118 et 152.)

[Basu 2012b] K. Basu, V. Debusschere et S. Bacha. *Load identification from power recordings at meter panel in residential households*. In Electrical Machines (ICEM), 2012 XXth International Conference on, pages 2098–2104, Sept 2012. (Referenced on page 2.)

[Basu 2012c] K. Basu, V. Debusschere et S. Bacha. *Load identification from power recordings at meter panel in residential households*. In Electrical Machines (ICEM), International Conference on, pages 2098–2104, 2-5 September 2012. (Referenced on page 163.)

[Basu 2013a] K. Basu, V. Debusschere et S. Bacha. *Residential appliance identification and future usage prediction from smart meter*. In Industrial Electronics Society, IECON 2013 - 39th Annual Conference of the IEEE, pages 4994–4999, Nov 2013. (Referenced on pages 164 et 171.)

[Basu 2013b] K. Basu, L. Hawarah, N. Arghira, H. Joumaa et S. Ploix. *A prediction system for home appliance usage*. Energy and Buildings, vol. 67, no. 1, pages 668–679, 2013. (Referenced on pages 7 et 94.)

[Basu 2014] K. Basu, V. Debusschere, S. Bacha, U. Maulik et S. Bondyopadhyay. *Non Intrusive Load Monitoring : A Temporal Multi-Label Classification Approach*. Industrial Informatics, IEEE Transactions on, vol. PP, no. 99, pages 1–1, 2014. (Referenced on page 2.)

[Berges 2010] M.E. Berges, E. Goldman, H.S. Matthews et L. Soibelman. *Enhancing Electricity Audits in Residential Buildings with Nonintrusive Load Monitoring*. Journal of Industrial Ecology, vol. 14, no. 5, pages 844–858, 2010. (Referenced on pages 11, 130 et 159.)

[Bier 2013] T. Bier, D. Benyoucef, D.O. Abdeslam, J. Merckle et P. Kleinf. *Smart meter systems measurements for the verification of the detection & classification algorithms*. In Industrial Electronics Society, IECON 2013-39th Annual Conference of the IEEE, pages 5000–5005. IEEE, 2013. (Referenced on pages 11 et 47.)

[Birt 2012] B.J. Birt, G.R. Newsham, I. Beausoleil-Morrison, M.M. Armstrong, N. Saldanha et I.H. Rowlands. *Disaggregating categories of electrical energy end-use from whole-house hourly data*. Energy and Buildings, vol. 50, pages 93–102, 2012. (Referenced on pages 10, 11 et 159.)

[Chawla 2011] N.V. Chawla, K.W. Bowyer, L.O. Hall et W.P. Kegelmeyer. *SMOTE : synthetic minority over-sampling technique*. arXiv preprint arXiv :1106.1813, 2011. (Referenced on page 68.)

[Chen 2001] B.J. Chen, M.W. Chang et C.J. Lin. *Load Forecasting Using Support Vector Machines : A Study on EUNITE Competition 2001*. Rapport technique, European Network on Intelligent TEchnologies for Smart Adaptive Systems, 2001. (Referenced on pages 18 et 92.)

[Clastres 2010] C. Clastres, T.T. Ha Pham, F. Wurtz et S. Bacha. *Ancillary services and optimal household energy management with photovoltaic production*. Energy, vol. 35, no. 1, pages 55–64, 2010. (Referenced on page 1.)

[Constanzo 2011] G.T. Constanzo. *Demand Side Management in the Smart Grid*. PhD thesis, Université de Montréal, Novembre 2011. (Referenced on page 132.)

[Debusschere 2014a] V. Debusschere, K. Basu et S. Bacha. *Identification et prédiction non intrusive de l'état des charges dans les bâtiments résidentiels à partir de mesures compteur à échantillonnage réduit*. In Symposium de Génie Électrique, 2014. (Referenced on page 3.)

[Debusschere 2014b] V. Debusschere, K. Basu et S. Bacha. *Identification non instrusive des charges dans un bâtiment résidentiel à partir de mesures compteur à faible pas de temps*. In Symposium de Genie Electrique, Cachan, 2014. (Referenced on page 3.)

[Debusschere 2014c] V. Debusschere, W. R. L. Garcia, K. Basu et S. Bacha. *Système de management énergétique résidentiel prédictif sous critères technico-économique*. In Symposium de Génie Électrique, 2014. (Referenced on page 3.)

[Debusschere 2014d] V. Debusschere, WRL. Garcia, K. Basu et S. Bacha. *Bilan sur cycle de vie des flux énergétiques dans les bâtiments résidentiels incluant de la production et du stockage*. In 3ème Conférence francophone sur l'éco-conception en Génie Electrique CONFREGE, 2014. (Referenced on page 3.)

[Dong 2012] M. Dong, P.C. Meira, W. Xu et W. Freitas. *An Event Window Based Load Monitoring Technique for Smart Meters*. Smart Grid, IEEE Transactions on, vol. 3, no. 2, pages 787–796, 2012. (Referenced on pages 50 et 160.)

[Douzal-Chouakria 2009] A. Douzal-Chouakria, A. Diallo et F. Giroud. *Adaptive clustering for time series : application for identifying cell cycle expressed genes*. Computational Statistics & Data Analysis, vol. 53, no. 4, pages 1414–1426, 2009. (Referenced on page 65.)

[Douzal-Chouakria 2012] A. Douzal-Chouakria et C. Amblard. *Classification trees for time series*. Pattern Recognition, vol. 45, no. 3, pages 1076–1091, 2012. (Referenced on page 63.)

[Eddy 1998] S.R. Eddy. *Profile hidden Markov models*. Bioinformatics, vol. 14, no. 9, pages 755–763, 1998. (Referenced on page 47.)

[EDF 2014] EDF. *Prix d'un kilowattheure en France*. EDF.fr. (2014). Retrieved 05 26, 2014, from Fournisseur electricite :, Accessed in septembre, 2014. (Referenced on page 139.)

[Elmahaiawy 2010] A.M. Elmahaiawy, N. Elfishawy et M.N. El-Dien. *Anticipation the consumed electrical power in smart home using evolutionary algorithms*. In Multimedia Computing and Information Technology (MCIT), 2010 International Conference on, pages 81–84. IEEE, 2010. (Referenced on page 15.)

[Engle 1992] R.F. Engle, C. Mustafa et J. Rice. *Modelling peak electricity demand*. Journal of Forecasting,, page 241–251, 1992. (Referenced on page 90.)

[Eynard 2010] J. Eynard. *Gestion optimale de l'énergie dans un procédé multi-source pour le chauffage de bâtiments*. PhD thesis, l'Université de Perpignan Via Domitia, 2010. (Referenced on page 19.)

[Farinaccio 1999] L. Farinaccio et R. Zmeureanu. *Using a pattern recognition approach to disaggregate the total electricity consumption in a house into the major end-uses*. Energy and Buildings, vol. 30, no. 3, pages 245–259, 1999. (Referenced on page 9.)

[Feinberg 2005] E.A. Feinberg et D. Genethliou. *Load Forecasting*. In Joe H. Chow, Felix F. Wu et James Momoh, editeurs, Applied Mathematics for Restructured Electric Power Systems, Power Electronics and Power Systems, pages 269–285. Springer US, 2005. (Referenced on pages 18 et 90.)

[Fernandes 2013] R.A.S. Fernandes, I.N. da Silva et M. Oleskovicz. *Load Profile Identification Interface for Consumer Online Monitoring Purposes in Smart Grids*. Industrial Informatics, IEEE Transactions on, vol. 9, pages 1507–1517, 2013. (Referenced on pages 11 et 159.)

[Figueiredo 2011] M.B Figueiredo, A. Almeida et B. Ribeiro. *An experimental study on electrical signature identification of non-intrusive load monitoring (NILM) systems*. In Proceedings of the 10th international conference on Adaptive and natural computing algorithms, Volume Part II, ICANNGA'11, pages 31–40, Berlin, Heidelberg, 2011. Springer-Verlag. (Referenced on page 159.)

[Ha 2006a] L.D. Ha, S. Ploix, E. Zamai et M. Jacomino. *Tabu Search for the optimization of household energy consumption*. In Information Reuse and Intergration (IRI) IEEE International Conference on, Waikoloa Hawaii, USA, September 16-18 2006. (Referenced on page 19.)

[Ha 2006b] S. Ha, H. Jung et Y. Oh. *Method to analyze user behavior in home environment.* Personal Ubiquitous Comput., vol. 10, pages 110–121, January 2006. (Referenced on page 15.)

[Ha 2008] L.D. Ha, S. Ploix, E Zamaï et M. Jacomino. *Realtimes dynamic optimization for demand-side load management.* International Journal of Management Science and Engineering Management, vol. 3, no. 4, November 2008. (Referenced on page 19.)

[Ha 2009] L.D. Ha, S. Ploix, F. Wurtz, P. Perichon et J. Merten. *Energy Management System for a photovoltaic grid-connected building.* In 24th EU PVSEC and 4$^{\text{th}}$ World Conference on Photovoltaic Energy Conversion, Hamburg, Germany, September, 21-26 2009. (Referenced on page 19.)

[Ha 2010] L.D. Ha, S. Ploix, M. Jacomino et H. Le Minh. Energy management, chapitre A mixed integer programming formulation of the home energy management problem. ISBN 978-953-307-065-0. INTECH, 2010. (Referenced on page 15.)

[Hart 1992] G.W. Hart. *Nonintrusive appliance load monitoring.* Proceedings of IEEE, vol. 80, no. 12, pages 1870–1891, 1992. (Referenced on pages 10, 11, 149, 159 et 160.)

[Hartigan 1979] J.A. Hartigan et M.A. Wong. *Algorithm AS 136 : A k-means clustering algorithm.* Applied statistics, pages 100–108, 1979. (Referenced on page 35.)

[Hawarah 2010] L. Hawarah, S. Ploix et M. Jacomino. *User Behavior Prediction in Energy Consumption in Housing Using Bayesian Networks.* In Artificial Intelligence and Soft Computing, volume 6113 of *Lecture Notes in Computer Science*, pages 372–379. Springer Berlin / Heidelberg, 2010. ploix2010d. (Referenced on pages 17 et 90.)

[Hecht-Nielsen 1989] R. Hecht-Nielsen. *Theory of the backpropagation neural network.* In Neural Networks, 1989. IJCNN., International Joint Conference on, pages 593–605. IEEE, 1989. (Referenced on page 92.)

[Henze 2003] G.P. Henze et R.H. Dodier. *Adaptive Optimal Control of a Grid-Independent Photovoltaic System.* Transactions of the ASME, vol. 125, pages 34–42, 2003. (Referenced on page 19.)

[Hippert 2001] H.S. Hippert, C.E. Pedreira et R.C. Souza. *Neural networks for short-term load forecasting : a review and evaluation.* Power Systems, IEEE Transactions on, vol. 16, no. 1, pages 44–55, feb 2001. (Referenced on pages 18 et 91.)

[Ho 1990] K.L. Ho, Y.Y. Hsu, C.F. Chen, T.E. Lee, C.C. Liang, T.S. Lai et K.K. Chen. *Short term load forecasting of Taiwan power system using a knowledge-based expert system.* Power Systems, IEEE Transactions on, vol. 5, no. 4, pages 1214–1221, nov 1990. (Referenced on page 91.)

[House 1995] J.M. House et T.F. Smith. *Optimal Control of Building and HVAC Systems.* In Proceddings of the American Control Conference, Seattle, Washington, 1995. (Referenced on page 19.)

[Humphries 1997] L.S. Humphries, J.D. Pritchett, G. Rasmussen et D.L. Voita. *Home automation system*, April 1997. US Patent 5,621,662. (Referenced on page 17.)

[Intille 2003] S.S. Intille, E.M. Tapia, J. Rondoni, J. Beaudin, C. Kukla, S. Agarwal, L. Bao et K. Larson. *Tools for studying behavior and technology in natural settings.* In UbiComp 2003 : Ubiquitous Computing, pages 157–174. Springer, 2003. (Referenced on page 21.)

[Intille 2005] S.S. Intille, K. Larson, J.S. Beaudin, J. Nawyn, E.M. Tapia et P. Kaushik. *A living laboratory for the design and evaluation of ubiquitous computing technologies.* In In Extended Abstracts of the 2005 Conference on Human Factors in Computing Systems, pages 1941–1944. ACM Press, 2005. (Referenced on page 20.)

[Intille 2006a] S.S. Intille. *The goal : smart people, not smart homes.* In International Conference on Smart Homes and Health Telematics, IOS. Press, 2006. (Referenced on page 20.)

[Intille 2006b] S.S. Intille, K. Larson, E. Tapia, J. Beaudin, P. Kaushik, J. Nawyn et R. Rockinson. *Using a Live-In Laboratory for Ubiquitous Computing Research.* Pervasive Computing, 2006. (Referenced on page 20.)

[Jolliffe 2005] I. Jolliffe. Principal component analysis. Wiley Online Library, 2005. (Referenced on page 36.)

[Kadous 2002] M.W. Kadous. *Temporal classification : Extending the classification paradigm to multivariate time series.* PhD thesis, The University of New South Wales, 2002. (Referenced on pages 50 et 160.)

[Kalogridis 2010] G. Kalogridis, C. Efthymiou, S.Z. Denic, T.A. Lewis et R. Cepeda. *Privacy for Smart Meters : Towards Undetectable Appliance Load Signatures.* In Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on, pages 232–237, oct. 2010. (Referenced on pages 11 et 160.)

[Khotanzad 1998] A. Khotanzad, R. Afkhami-Rohani et D. Maratukulam. *ANNSTLF-Artificial Neural Network Short-Term Load Forecaster generation three.* Power Systems, IEEE Transactions on, vol. 13, no. 4, pages 1413–1422, nov 1998. (Referenced on pages 18 et 91.)

[Kohavi 1995] R. Kohavi. *The Power of Decision Tables.* In Proceedings of the European Conference on Machine Learning, pages 174–189. Springer Verlag, 1995. (Referenced on page 92.)

[Kohavi 1998] R. Kohavi et F. Provost. *Glossary of Terms.* Editorial for the Special Issue on Applications of Machine Learning and the Knowledge Discovery Process, vol. 30, pages 271–274, February 1998. (Referenced on page 68.)

[Kolter 2010] J.Z. Kolter, S. Batra et A.Y. Ng. *Energy disaggregation via disciminative sparce coding.* Advances in Neural Information Processing Systems, vol. 23, no. 7, pages 1153–1161, 2010. (Referenced on pages 11, 47 et 160.)

[Kouveletsou 2012] M. Kouveletsou, N. Sakkas, S. Garvin, M. Batic, D. Reccardo et R. Sterling. *Simulating energy use and energy pricing in buildings The case of electricity.* Energy and Buildings, vol. 54, no. 0, pages 96–104, 2012. (Referenced on page 130.)

[Labeeuw 2013] W. Labeeuw et G. Deconinck. *Residential Electrical Load Model Based on Mixture Model Clustering and Markov Models.* Industrial Informatics, IEEE Transactions on, vol. 9, no. 3, pages 1561–1569, 2013. (Referenced on pages 50 et 160.)

[Lammarsch 2009] T. Lammarsch, W. Aigner, A Bertone, J. Gartner, E. Mayr, S. Miksch et M. Smuc. *Hierarchical Temporal Patterns and Interactive Aggregated Views for Pixel-Based Visualizations.* In Information Visualisation, 2009 13th International Conference, pages 44–50, July 2009. (Referenced on pages 29 et 30.)

[Laxman 2006] S. Laxman et P.S. Sastry. *A survey of temporal data mining.* Sadhana, vol. 31, no. 2, pages 173–198, 2006. (Referenced on page 159.)

[Les Energies Renouvelables 2011] Les Energies Renouvelables. *Tarif de rachat electricite photovoltaique pour 2014*, Accessed in septembre, 2014 2011. (Referenced on page 139.)

[Li 2013] J. Li et N. Allinson. *Building Recognition Using Local Oriented Features.* Industrial Informatics, IEEE Transactions on, vol. 9, no. 3, pages 1697–1704, 2013. (Referenced on page 11.)

[Livengood 2009] R.L.D. Livengood. *The Energy Box : Locally Automated Optimal Control of Residential Electricity Usage.* Rapport technique, Center for Engineering Systems Fundamentals, Engineering Systems Division, MIT, Cambridge, Massachusetts, 2009. (Referenced on pages 130 et 152.)

[Long 2005] D.H. Long, S. Ploix, E. Zamaï et M. Jacomino. *Control of energy consumption in home automation by ressource constraint scheduling.* In The 15th international conference on control systems and computer science, Buccarest, Romania, 2005. (Referenced on page 19.)

[Miller 2006] C.H. Miller. *Patent US2006/0276938 : Optimized Energy Management System*, 2006. (Referenced on page 20.)

[Missaoui 2010] R. Missaoui, G. Warkozek, S. Abras, S. Ploix et S. Bacha. *Simulation temps réel pour la gestion des flux énergétiques dans l'habitat.* In Conférence de l'International Building Performance Simulation Association (IBPSA), Moret-sur-Loing, France, November 9-10 2010. (Referenced on page 19.)

[Missaoui 2014] R. Missaoui, H. Joumaa, S. Ploix et S. Bacha. *Managing energy smart homes according to energy prices : analysis of a building energy management system.* Energy and Buildings, vol. 71, pages 155–167, 2014. (Referenced on page 138.)

[Muselli 2000] M. Muselli, G. Notton, P. Poggi et A. Louche. *PV-hybrid power system sizing incorporating battery storage : an analysis via simulation calculations.* Renewable Energy, vol. 20, pages 1–7, 2000. (Referenced on page 19.)

[Negenborn 2007] R.R. Negenborn. *Multi-agent model predictive control with applications to power networks.* PhD thesis, Technische Universiteit Delft, Delft, Netherlands, december 2007. (Referenced on page 19.)

[Norford 1996] L.K. Norford et S.B. Leeb. *Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algorithms.* Energy and Buildings, vol. 24, no. 1, pages 51–64, 1996. (Referenced on pages 11 et 160.)

[Observatoire Electricité 2014] Observatoire Electricité. *La consommation d'énergie en France en 2012.* Rapport technique, Union Française de l'Electricité, 2013, Accessed in September 2014. (Referenced on page 1.)

[Onoda 2000] T. Onoda, G. Rätsch et K.R. Müller. *Applying Support Vector Machines and Boosting to a Non-Intrusive Monitoring System for Household Electric Appliances with Inverters*, 2000. (Referenced on pages 62 et 160.)

[Palensky 1997] P. Palensky et R. Posta. *Demand side management in private home using LONWORKS.* In Factory Communication Systems, IEEE International Workshop on, 1997. (Referenced on page 16.)

[Palensky 2011] P. Palensky et D. Dietrich. *Demand side management : Demand response, intelligent energy systems, and smart loads.* Industrial Informatics, IEEE Transactions on, vol. 7, no. 3, pages 381–388, 2011. (Referenced on pages 7, 130 et 159.)

[Park 1991] D.C. Park, M.A. El-Sharkawi, R.J. Marks, L.E. Atlas et M.J. Damborg. *Electric load forecasting using an artificial neural network.* Power Systems, IEEE Transactions on, vol. 6, no. 2, pages 442–449, may 1991. (Referenced on pages 18 et 91.)

[Parson 2011] O. Parson, S. Ghosh, M. Weal et A. Rogers. *Using Hidden Markov Models for Iterative Non-intrusive Appliance Monitoring.* In Neural Information Processing Systems workshop on Machine Learning for Sustainability, December 2011. Event Dates : 17 December 2011. (Referenced on pages 12 et 50.)

[Pearl 1986] J. Pearl. *Fusion, propagation, and structuring in belief networks.* Artif. Intell., vol. 29, no. 3, pages 241–288, 1986. (Referenced on page 93.)

[Pedersen 2008] L. Pedersen, J. Stang et R. Ulseth. *Load prediction method for heat and electricity demand in buildings for the purpose of planning for mixed energy distribution systems.* Energy and Buildings, vol. 40, no. 7, pages 1124–1134, 2008. (Referenced on page 19.)

[Ploix 2010] S. Ploix, Ha Duy Long, S. Abras, M. Jacomino, D. Wurtz et S. Bacha. *patent US/EU 61/326,542 : Environment and method for managing services in a living place*, 2010. (Referenced on page 20.)

[Project 2009] SmartHouse/SmartGrid Project. *Deliverable 1.2 - Technology trends for SmartHouse/SmartGrid.* Rapport technique, EU/ICT Project EU FP7-ICT-2007-2 STREP 224628, 2009. (Referenced on page 20.)

[Prudenzi 2002] A. Prudenzi. *A neuron nets based procedure for identifying domestic appliances pattern-of-use from energy recordings at meter panel.* In Power Engineering Society Winter Meeting, 2002. IEEE, volume 2, pages 941–946, 2002. (Referenced on page 11.)

[Quinlan 1986] J.R. Quinlan. *Induction of decision trees.* Machine Learning, vol. 1, pages 81–106, 1986. (Referenced on page 60.)

[Rabiner 1986] L. Rabiner et B.H. Juang. *An introduction to hidden Markov models.* ASSP Magazine, IEEE, vol. 3, no. 1, pages 4–16, 1986. (Referenced on page 66.)

[Read 2011] J. Read, B. Pfahringer, G. Holmes et E. Frank. *Classifier chains for multi-label classification.* Machine Learning, vol. 85, pages 333–359, 2011. (Referenced on page 75.)

[Riffoneau 2009] Y. Riffoneau. *Gestion des Flux Energétiques dans un Systeme Photovoltaïque avec Stockage Connecté au Réseau.* PhD thesis, Grenoble : Université Joseph Fourier, G2Elab, EEATS, Octobre 2009. (Referenced on pages 131, 132, 135 et 152.)

[Riffoneau 2010] Y. Riffoneau. *Gestion des flux énergétiques dans un système photovoltaïque avec stockage connecté au réseau.* PhD thesis, Grenoble Institute of Technology (INPG), 2010. (Referenced on pages 16 et 19.)

[Rigo-Mariani 2014] R. Rigo-Mariani, B. Sareni, X. Roboam et Ch. Turpin. *Optimal power dispatching strategies in smart-microgrids with storage.* Renewable and Sustainable Energy Reviews, vol. 40, pages 649–658, 2014. (Referenced on page 19.)

[Rte 2011] Rte. *Consommation Française d'électricité : Caractéristiques et méthode de prévision.* Rapport technique, Réseau de transport d'électricité, 2011. (Referenced on page 18.)

[Rte 2014] Rte. *Courbes journalières de consommation.* http ://clients.rte-france.com/lang/fr/visiteurs/vie/courbes.jsp, Accessed on September 2014. Réseau de transport d'électricité. (Referenced on page 18.)

[Siano 2014] P. Siano. *Demand response and smart grids - A survey.* Renewable and Sustainable Energy Reviews, vol. 30, no. 1, pages 461–478, 2014. (Referenced on pages 7, 130 et 159.)

[Strasser 2013] T. Strasser, F. Andrén, M. Merdan et A. Prostejovsky. *Review of Trends and Challenges in Smart Grids : An Automation Point of View.* In Industrial Applications of Holonic and Multi-Agent Systems, volume 8062 of *Lecture Notes in Computer Science*, pages 1–12. Springer Berlin Heidelberg, 2013. (Referenced on page 7.)

[Stum 1997] K. Stum, R. Mosier et T. Haasl. *Energy Management Systems.* Rapport technique, Porland Energy Conservation Inc.(PECI), 1997. (Referenced on page 19.)

[Tapia 2004] E.M. Tapia, S.S. Intille et K. Larson. *Activity Recognition in the Home using Simple and Ubiquitous Sensors.* In Pervasive Computing, pages 158–175, 2004. (Referenced on page 21.)

[Thiaux 2010] Y. Thiaux. *Optimisation des profils de consommation pour minimiser les coûts économique et énergétique sur cycle de vie des systèmes photovoltaïques autonomes et hybrides évaluation du potentiel de la téchnologie Li-ion.* PhD thesis, Ecole Normale Supérieure de Cachan, 2010. (Referenced on page 133.)

[Tran 2012] V.G. Tran, V. Debusschere et S. Bacha. *Hourly server workload forecasting up to 168 hours ahead using Seasonal ARIMA model.* In Industrial Technology (ICIT), 2012 IEEE International Conference on, pages 1127–1131. IEEE, 2012. (Referenced on page 90.)

[Tran 2013] V.G. Tran, V. Debusschere et S. Bacha. *Neural networks for web server workload forecasting*. In Industrial Technology (ICIT), 2013 IEEE International Conference on, pages 1152–1156. IEEE, 2013. (Referenced on page 90.)

[Tsoumakas 2007] G. Tsoumakas et I. Katakis. *Multi Label Classification : An Overview*. International Journal of Data Warehousing and Mining, vol. 3, no. 3, pages 1–13, 2007. (Referenced on pages 58, 59, 69 et 75.)

[Tsoumakas 2010] G. Tsoumakas, I. Katakis et I. Vlahavas. *Mining Multi-label Data*. In Oded Maimon et Lior Rokach, editeurs, Data Mining and Knowledge Discovery Handbook, chapitre 34, pages 667–685. Springer US, Boston, MA, 2010. (Referenced on pages 60, 74 et 159.)

[Tsoumakas 2011] G. Tsoumakas, E. Spyromitros-Xioufis, J. Vilcek et I. Vlahavas. *Mulan : A Java Library for Multi-Label Learning*. Journal of Machine Learning Research, vol. 12, pages 2411–2414, 2011. (Referenced on pages 58, 59 et 69.)

[Wacks 1991] K.P. Wacks. *Utility load management using home automation*. Consumer Electronics, IEEE Transactions on, vol. 37, no. 2, pages 168–174, 1991. (Referenced on page 130.)

[Wacks 1993] K.P. Wacks. *The impact of home automation on power electronics*. In Applied Power Electronics Conference and Exposition, pages 3–9, 1993. (Referenced on page 16.)

[Warkozek 2009] G. Warkozek, M. Jacomino, S. Ploix et F. Wurtz. *Generic formulation of optimization problems for energy management : solving difficulties, practical and mathematical analysis*. In Proceedings of the 8^{th} International Symposium on Electric and Magnetic Fields (EMF2009), Mondovi, Italy, May 26th-29th 2009. (Referenced on page 20.)

[Weka 2014] Machine Learning Project Weka. *Weka*, Accessed in September 2014. (Referenced on pages 61, 62 et 98.)

[Witten 2011] I.H. Witten, E. Frank et M.A. Hall. Data mining : Practical machine learning tools and techniques. Morgan Kaufmann series in data management systems. Elsevier Science & Technology, 2011. (Referenced on page 93.)

[Wu 2007] X. Wu, V. Kumar, J.R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G.J. McLachlan, A. Ng, B. Liu, P.S. Yu, Z.H. Zhou, M. Steinbach, D.J. Hand et D. Steinberg. *Top 10 algorithms in data mining*. Knowl. Inf. Syst., vol. 14, pages 1–37, December 2007. (Referenced on page 75.)

[Zeifman 2011] M. Zeifman et R. Kurt. *Nonintrusive appliance load monitoring : Review and outlook*. Consumer Electronics, IEEE Transactions on, vol. 57, no. 1, pages 76–84, 2011. (Referenced on pages 11 et 160.)

[Zhang 2007] M. Zhang et Z. Zhou. *ML-KNN : A lazy learning approach to multi-label learning*. Pattern Recognition, vol. 40, no. 7, pages 2038–2048, 2007. (Referenced on pages 65 et 75.)

[Zhou 2005] G. Zhou et M. Krarti. *Parametric Analysis of Active and Passive Building Thermal Storage Utilization*. Journal of Solar Energy Engineering, vol. 127, pages 37–46, 2005. (Referenced on page 19.)

Annexe A

Annexure

A.1 Sample dataset

In the figure A.1 and A.2 a screen shot of the raw dataset is presented. In the figure A.3 the dataset after feature extraction is presented.

	date	Fridge (Kitchen, 180l)	Halogen lamp 1 (0,5kW)	Power supply for wood boiler ()	Site consumption ()	Total site light consumption ()	Vertical freezer (Cellar, 145l)	Washing
2	Thursday, January 22, 1998 2:00 PM	11.0	0.0	64.0	599.0	0.0	34.0	0.0
3	Thursday, January 22, 1998 3:00 PM	10.0	0.0	65.0	666.0	108.0	32.0	0.0
4	Thursday, January 22, 1998 4:00 PM	21.0	0.0	65.0	550.0	70.0	31.0	0.0
5	Thursday, January 22, 1998 5:00 PM	10.0	45.0	64.0	714.0	124.0	31.0	0.0
6	Thursday, January 22, 1998 6:00 PM	19.0	376.0	60.0	1644.0	192.0	31.0	0.0
7	Thursday, January 22, 1998 7:00 PM	19.0	278.0	62.0	999.0	103.0	39.0	0.0
8	Thursday, January 22, 1998 8:00 PM	18.0	106.0	62.0	797.0	166.0	44.0	0.0
9	Thursday, January 22, 1998 9:00 PM	18.0	0.0	64.0	514.0	38.0	40.0	0.0
10	Thursday, January 22, 1998 10:00 PM	18.0	0.0	64.0	242.0	0.0	29.0	0.0
11	Thursday, January 22, 1998 11:00 PM	10.0	0.0	66.0	220.0	0.0	30.0	0.0
12	Friday, January 23, 1998 12:00 AM	20.0	0.0	66.0	226.0	0.0	30.0	0.0
13	Friday, January 23, 1998 1:00 AM	19.0	0.0	65.0	198.0	0.0	30.0	0.0
14	Friday, January 23, 1998 2:00 AM	11.0	0.0	65.0	217.0	0.0	30.0	0.0
15	Friday, January 23, 1998 3:00 AM	14.0	0.0	67.0	224.0	0.0	30.0	0.0
16	Friday, January 23, 1998 4:00 AM	17.0	0.0	66.0	192.0	0.0	31.0	0.0
17	Friday, January 23, 1998 5:00 AM	11.0	0.0	66.0	224.0	0.0	39.0	0.0
18	Friday, January 23, 1998 6:00 AM	12.0	0.0	65.0	276.0	47.0	38.0	0.0
19	Friday, January 23, 1998 7:00 AM	22.0	0.0	63.0	434.0	143.0	35.0	0.0
20	Friday, January 23, 1998 8:00 AM	12.0	0.0	65.0	249.0	80.0	29.0	0.0
21	Friday, January 23, 1998 9:00 AM	11.0	0.0	66.0	564.0	3.0	30.0	0.0
22	Friday, January 23, 1998 10:00 AM	14.0	0.0	66.0	220.0	18.0	30.0	0.0
23	Friday, January 23, 1998 11:00 AM	15.0	0.0	65.0	161.0	0.0	31.0	0.0
24	Friday, January 23, 1998 12:00 PM	15.0	0.0	65.0	190.0	34.0	31.0	0.0
25	Friday, January 23, 1998 1:00 PM	23.0	0.0	63.0	208.0	37.0	30.0	0.0
26	Friday, January 23, 1998 2:00 PM	13.0	0.0	64.0	157.0	0.0	32.0	0.0
27	Friday, January 23, 1998 3:00 PM	12.0	0.0	64.0	181.0	14.0	43.0	0.0
28	Friday, January 23, 1998 4:00 PM	12.0	0.0	65.0	168.0	3.0	42.0	0.0

Figure A.1 – An example of the raw dataset screenshot

date	Microwave oven	Vertical freezer	Site consumpti	Dish washer (1	Electric oven ()
04/28/1999 12:00	50	65	666	0	0
04/28/1999 13:00	44	53	486	0	0
04/28/1999 14:00	0	72	452	0	0
04/28/1999 15:00	19	57	431	0	0
04/28/1999 16:00	26	64	397	0	0
04/28/1999 17:00	0	66	413	0	0
04/28/1999 18:00	0	54	352	0	0
04/28/1999 19:00	0	73	800	403	0
04/28/1999 20:00	0	52	1862	1301	0
04/28/1999 21:00	0	58	535	0	0
04/28/1999 22:00	0	70	450	0	0
04/28/1999 23:00	0	55	376	0	0
04/29/1999 00:00	0	54	347	0	0
04/29/1999 01:00	0	69	345	0	0
04/29/1999 02:00	0	62	348	0	0
04/29/1999 03:00	0	54	334	0	0
04/29/1999 04:00	0	55	344	0	0
04/29/1999 05:00	0	67	353	0	0
04/29/1999 06:00	100	62	669	0	0
04/29/1999 07:00	32	54	537	0	0
04/29/1999 08:00	48	54	426	0	0
04/29/1999 09:00	0	75	364	0	0
04/29/1999 10:00	0	58	367	0	0
04/29/1999 11:00	0	71	1395	0	964
04/29/1999 12:00	124	59	956	0	284

Figure A.2 – An example of the raw dataset screenshot

A.2 Database groping (parallel coordinate plot)

In the figure A.4 the parallel coordinate plot for the houses in IRESE database is shown from the features extracted in chapter 4. A.5 the cluster centre are also represented in parallel coordinates of the feature. The coordinates are normalised between 0-1.

Hour	Day	Months	Season	ConsH-24	ConsH-23	ConsH-22	ConsH-21	ConsH-20	ConsH-19	ConsH-18
12	Thursday	4	1	0	0	0	0	0	0	0
13	Thursday	4	1	0	0	0	0	0	0	0
14	Thursday	4	1	0	0	0	0	0	0	0
15	Thursday	4	1	0	0	0	0	0	0	0
16	Thursday	4	1	0	0	0	0	0	0	0
17	Thursday	4	1	0	0	0	0	0	0	0
18	Thursday	4	1	0	0	0	0	0	0	0
19	Thursday	4	1	0	0	0	0	0	0	0
20	Thursday	4	1	0	0	0	0	0	0	0
21	Thursday	4	1	0	0	0	0	0	0	0
22	Thursday	4	1	0	0	0	0	0	0	0
23	Thursday	4	1	0	0	0	0	0	0	0
0	Friday	4	1	0	0	0	0	0	0	0
1	Friday	4	1	0	0	0	0	0	0	0
2	Friday	4	1	0	0	0	0	0	0	0
3	Friday	4	1	0	0	0	0	0	0	0
4	Friday	4	1	0	0	0	0	0	0	0
5	Friday	4	1	0	0	0	0	0	0	1
6	Friday	4	1	0	0	0	0	0	1	1
7	Friday	4	1	0	0	0	0	1	1	0
8	Friday	4	1	0	0	0	1	1	0	0
9	Friday	4	1	0	0	1	1	0	0	0

FIGURE A.3 – An example of the raw dataset screenshot after feature extration

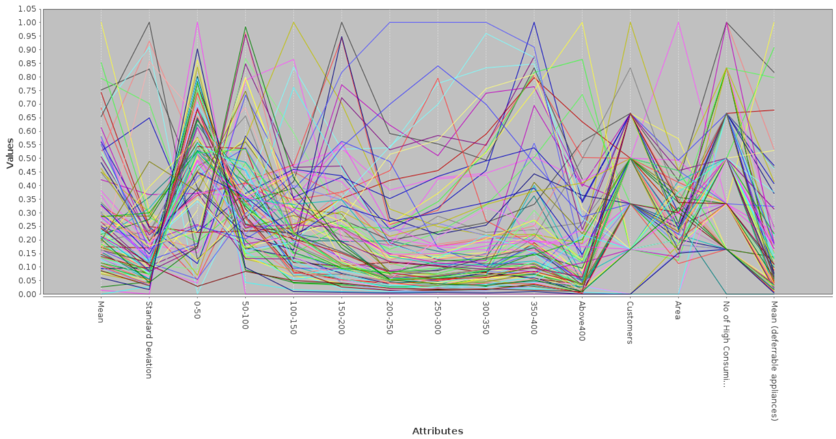


FIGURE A.4 – Parallel coordinate plot for the IRESE dataset (All the houses)

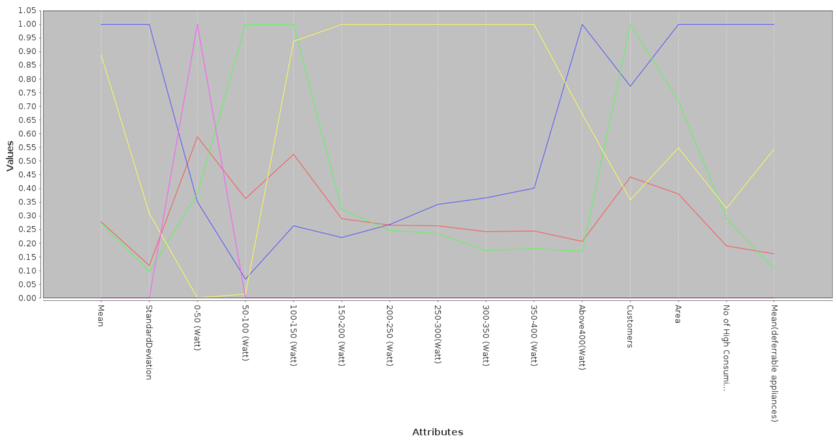


FIGURE A.5 – Parallel coordinate plot for the IRESE dataset (cluster centre)

A.3 Load identification frequency analysis

In the figure A.6 the load identification of the Appliance water heater is compared with actual case. In is a hourly frequency plot to observe the number of times the appliance consumed in this period. The results show that the nature of both the curves in similar.

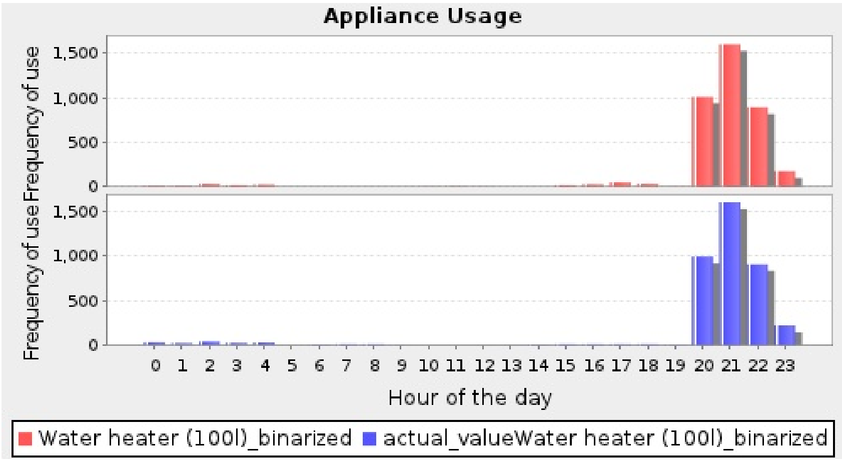


FIGURE A.6 – The water heater usage identification is shown

Résumé — french version Nous abordons dans ces travaux l'identification non intrusive des charges des bâtiments résidentiels ainsi que la prédiction de leur état futur. L'originalité de ces travaux réside dans la méthode utilisée pour obtenir les résultats voulus, à savoir l'analyse statistique des données (algorithmes de classification). Celle-ci se base sur des hypothèses réalistes et restrictives sans pour autant avoir de limitation sur les modèles comportementaux des charges (variations de charges ou modèles) ni besoin de la connaissance des changements d'état des charges. Ainsi, nous sommes en mesure d'identifier et/ou de prédire l'état des charges consommatrices d'énergie (et potentiellement contrôlables) en se basant uniquement sur une phase d'entrainement réduite et des mesures de puissance active agrégée sur un pas de mesure de dix minutes, préservant donc la vie privée des habitants. Dans cette communication, après avoir décrit la méthodologie développée pour classifier les charges et leurs états, ainsi que les connaissances métier fournies aux algorithmes, nous comparons les résultats d'identification pour cinq algorithmes tirés de l'état de l'art et les utilisons comme support d'application à la prédiction. Les algorithmes utilisés se différentient par leur capacité à traiter des problèmes plus ou moins complexe (notamment la prise en compte de relations entre les charges) et se ne révèlent pas tous appropriés à tout type de charge dans le bâtiment résidentiel.

Mots clés : Identification non intrusive, prédiction, bâtiment résidentiel, classification multi-étiquette, analyse de données.

English Version

Abstract — Smart metering is one of the fundamental units of a smart grid, as many further applications depend on the availability of fine-grained information of energy consumption and production. Demand response techniques can be substantially improved by processing smart meter data to extract relevant knowledge of appliances within a residence. The thesis aims at finding generic solutions for the non-intrusive load monitoring and future usage prediction of residential loads at a low sampling rate. Load monitoring refers to the dis-aggregation of individual loads from the total consumption at the smart meter. Future usage prediction of appliances are important from the energy management point of view. In this work, state of the art multi-label temporal classification techniques are implemented using novel set of features. Moreover, multi-label classifiers are able to take inter-appliance correlation into account. The methods are validated using a dataset of residential loads in 100 houses monitored over a duration of 1-year.

Keywords : Non-intrusive load monitoring, Appliance usage prediction, smart-meter, smart-grid, load monitoring, optimization, time series, machine learning, temporal classification, clustering

G2E-lab, 11 rue des Mathématiques, Grenoble Campus BP 46,
F-38402 Saint Martin d'Hères CEDEX